Topics in Bayesian Inference applied to Probabilistic Power Flow Analysis

Carlos David Zuluaga R.

Universidad Tecnológica de Pereira

This dissertation is submitted for the degree of

Doctor in Engineering

Research in Automatics

July 2018

To the memory of Valentina Zuluaga Ríos.

Declaration

I hereby declare that this thesis entitled "Topics in Bayesian Inference applied to Probabilistic Power Flow Analysis" was carried out by me for the degree of Doctor in Engineering under the guidance and supervision of Mauricio A. Álvarez, PhD and Lecturer in Machine Learning at the Department of Computer Science at the University of Sheffield.

The contents of this document are original and have not been submitted in whole or in part for consideration for any other degree or qualification in this, or any other university. This thesis is my own work and contains nothing which is the outcome of work done in collaboration with others, except as specified in the text and Acknowledgements.

Carlos David Zuluaga R.
July 2018

Acknowledgements

Foremost, I wish to thank my supervisor Mauricio A. Álvarez L. for his support, motivation and patience. I must admit that this thesis would not have been possible without his guidance and help.

I want to thank to the members of the Research Group in Automatics that have contributed to my personal and professional time at the Universidad Tecnológica de Pereira. I am also grateful with Edward Meeds and Richard Wilkinson, their suggestions were very helpful to improve the quality of this study.

I gratefully acknowledge the funding sources that made my Ph.D. work possible. I was funded by Department of Science, Technology and Innovation, Colciencias, Colombia, under the program: "programa nacional de formación de investigadores - Convocatoria 567 de 2012". This project has been partially funded by Colciencias under the research project 111074558696. My work was also supported by Vicerrectoría de Investigaciones, Innovación y Extensión, Universidad Tecnológica de Pereira, Colombia, under the program: Convocatoria interna para proyectos de investigación de estudiantes de Posgrado − Año 2015.

Lastly, I would like to thank my family for all their love and encouragement. Thank you.

Abstract

A probabilistic power flow (PPF) study is an essential tool for the analysis and planning of a power system when specific variables are considered as random variables with particular probability distributions. The most widely used method for solving the PPF problem is Monte Carlo simulation (MCS). Although MCS is accurate for obtaining the uncertainty of the state variables, it is also computationally expensive, since it relies on repetitive deterministic power flow solutions. On the other hand, MCS does not take into account the fact that previous knowledge of state variables might be available in terms of probability distributions.

In this thesis, we frame the PPF as a probabilistic inference problem, and instead of repetitively solving optimization problems, we use Bayesian inference for computing posterior distributions over state variables. We specifically use prior distributions for the state variables, and a likelihood function that relates the observations to the state variables. We apply Bayes theorem to obtain the posterior distribution over the state variables. By using a Bayesian inference perspective, we can model the state variables as random variables, and we do not need to solve heavy computational optimization methods for computing posterior distributions over state variables.

Here, we explore two perspectives for the PPF problem. As a first model, we use a hierarchical Bayesian model for the PPF analysis, specifying prior distributions over the state variables and a likelihood function that relates the state variables and the observations. In this first model, we use likelihood-based Bayesian methods, such as the Markov chain Monte Carlo and Hamiltonian Monte Carlo methods, for estimating the posterior distributions over state variables. As a second Bayesian perspective, we provide a likelihood-free Bayesian approach based on the Approximate Bayesian Computation philosophy, that incorporates the Jacobian computed from the power flow equations. Results in three different test systems show that the proposed methodologies are competitive alternatives for solving the PPF problem, and in some cases, they allow for reduction in computation time when compared to MCS.

Contents

List of Figures

List of Tables

Notation

Generalities

N	number of power system nodes
$N_{\mathbf{x}}$	number of state variables
N_s	number of observations or samples
$\mathbb{R}$	set of real numbers
ϵ	tolerance threshold in approximate Bayesian computation methods

Functions

$\mathbf{g}(\mathbf{x})$, $\mathbf{h}(\mathbf{x})$	nonlinear power flow equations
$H(\boldsymbol{\rho}, \mathbf{x})$	Hamiltonian function
$K(\boldsymbol{\rho})$	kinetic energy in $H(\boldsymbol{\rho}, \mathbf{x})$
$U(\mathbf{x})$	Potential energy in $H(\boldsymbol{\rho}, \mathbf{x})$
$p(z)$	probability distribution for z
$p(z \mid y)$	conditional probability of z given y
$q(x^* \mid x)$	proposal distribution of the new state x^* given a current state x
$\mathcal{N}(\mu, \sigma^2)$	Gaussian distribution with parameters μ and σ^2
$\mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$	multivariate Gaussian distribution with parameters $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$
$\mathcal{S}(\mathbf{x} \mid \boldsymbol{\mu}, \boldsymbol{\Lambda}, \nu)$	multivariate t distribution with parameters $\boldsymbol{\mu}$, $\boldsymbol{\Lambda}$ and ν
$\mathrm{Dis}(\boldsymbol{\mu}, \boldsymbol{\pi})$	discrete distribution with parameters $\boldsymbol{\mu}$ and $\boldsymbol{\pi}$
$\mathcal{U}(a, b)$	uniform distribution with parameters a and b
$s(\mathcal{D})$	summary statistics of the observed data
$\gamma_k^2(P, Q)$	metric between the probability measures P and Q
$k(x_i, x_j)$	dot product between the observation x_i and observation x_j

Vectors and matrices

$\mathbf{b}$	vector with entries given by the net active and reactive powers injected
$\mathbf{x}$	state variables: angles and voltages
$\mathbf{J}$	Jacobian matrix of $\mathbf{g}(\mathbf{x})$
$\mathbf{J}_i$	Jacobian matrix evaluated at $\mathbf{x}_i$
$\mathbf{\Sigma}$	covariance matrix
$\mathbf{\Lambda}$	precision matrix
$\boldsymbol{\pi}$	vector of probabilities for the components $(\mu_1, \mu_2, \ldots)$ in a discrete distribution
$\mathcal{D}$	observed data
$\mathcal{D}'$	simulated data

Abbreviations

ABC	Approximate Bayesian Computation
ABC MCMC	ABC based on a Markov Chain Monte Carlo approach
ABC SMC	ABC using a sequential Monte Carlo approach
BD	Bhattacharyya distance
DPF	Deterministic Power Flow
HBM	Hierarchical Bayesian Modeling
HMC	Hamiltonian Monte Carlo
JABC	Approximate Bayesian Computation method that incorporates the Jacobian matrix
K2ABC	Approximate Bayesian Computation method based on kernel embeddings (two kernels) and empirical distributions
MCMC	Markov Chain Monte Carlo
MCS	Monte Carlo simulation
MMD	Maximum Mean Discrepancy
PABC	Approximate Bayesian Computation method using Kernel Embeddings and a Parzen-Based Distance
PPF	Probabilistic power flow
PQ node	Load node
PV node	Voltage-controlled node
RE	Relative error

Chapter 1

Introduction

Probabilistic power flow (PPF) analysis is an essential tool for the analysis, operation and planning of a power system (PS) when the power generation and loads are considered as random variables with particular probability distributions. The PPF goal is to obtain probability distributions over voltages, angles and power flows between lines (Morales et al., 2010).

The most widely used method for solving the PPF problem is Monte Carlo simulation (MCS) (Carmona-Delgado et al., 2015; Gallego and Padilha-Feltrin, 2012; Le et al., 2013). MCS is based on statistical sampling of input random variables (loads, power generation and voltages at PV nodes) and the propagation of these samples through repetitive deterministic solutions of the PPF model.

The process described above has its shortcomings. On one hand, it does not take into account the fact that previous knowledge of state variables (angles at PQ and PV nodes, and voltages at PQ nodes) might be available in terms of probability distributions, for example: from normal operating conditions of a PS, the magnitude of the voltage variables is close to one per unit, therefore, we can define a specific probability distribution for these variables. That is, MCS does not consider the state variables as random variables within the PPF problem before observing different configurations of the input random variables. On the other hand, this sampling approach is computationally expensive, since it relies on repetitive deterministic power flow (DPF) solutions (Soleimanpour and Mohammadi, 2013; Zhang and Li, 2010). According to Soleimanpour and Mohammadi (2013), MCS took 14 seconds to analyze the IEEE 39-bus test system.

We address these shortcomings by formulating the PPF problem as a Bayesian inference problem. Bayesian inference requires the specification of prior distributions for the state variables, and a likelihood function that relates the observations to the state variables. We apply Bayes theorem to obtain the posterior distribution over the state variables. By including prior distributions over the state variables into the PPF problem, it will be possible to exploit an additional source of information that has not been used in simulation methods before, like MCS.

From the above, we are interested in exploring several open questions related to *how to incorporate the Bayesian inference framework into the PPF analysis*.

First, how to formulate the PPF problem under a Bayesian inference perspective, that is, *how to define prior distributions over the state variables* and *how to formulate or propose a likelihood function that relates the power injected and the state variables*. As a consequence of the Bayesian approach, the quantification of uncertainty over state variables is incorporated into the PPF problem, which is a novel contribution to the best of our knowledge. In the Bayesian paradigm, the likelihood function plays a central role. With this function, we can compute the uncertainty of the state variables after observing the input random variables, such as active and reactive powers injected. If these input variables are modeled by Gaussian distributions, we can define a likelihood function as a multivariate Gaussian distribution (see Dopazo et al. (1975)). However, if we consider the integration of the distributed generation to the PPF analysis, we can not use a simple multivariate Gaussian distribution, due to this type of power generation is necessary to use non-Gaussian distributions to model the uncertainty into the source of non-conventional energy. For example, it is common to use a Weibull distribution to model the uncertainty in the wind speed for wind energy or to employ a beta distribution to consider the solar radiation for solar energy (Aien et al., 2014; Soleimanpour and Mohammadi, 2013). In this case, we should define a specific likelihood function to introduce this type of generation into the Bayesian PPF problem.

After having specified the prior distributions over the state variables and the likelihood function, we can solve the Bayesian PPF problem by using likelihood-based methods. The most popular methods are Markov chain Monte Carlo (MCMC) (Murphy, 2012, pp. 438) and Hamiltonian Monte Carlo (HMC) (Bishop, 2006, pp. 548). The basic idea behind MCMC and HMC is to construct a Markov chain on the state space

whose stationary distribution is the posterior distribution over the state variables.

In a classical Bayesian inference problem, we would know beforehand the likelihood function. Since likelihood functions for PPF problems have not been discussed properly in the literature, we discuss alternatives for likelihood functions. Instead of choosing a particular form for the likelihood term in our Bayesian analysis, we will also use likelihood-free methods for obtaining the posterior distribution. Therefore, *we would like to know what likelihood-free Bayesian methods we can apply to solve this Bayesian PPF problem with a minor computational cost.* Likelihood-free methods are used to approximate the posterior distribution using simulation rather than likelihood calculations. These likelihood-free Bayesian approaches are commonly named as Approximate Bayesian computation (ABC) methods. In contrast to the MCS, the ABC methods employ distances between summary statistics of the simulated data and summary statistics of the observed data for accepting samples from the approximated posterior distribution. For the PPF problem, the observed data is composed of the active and reactive powers injected. Using this procedure, ABC methods can take less computation time for obtaining the posterior distribution of the state variables than MCS. However, in ABC methods, a crucial topic is how to choose the appropriate summary statistics.

1.1 Why Bayesian

From the Bayesian inference, we can do three important task: point estimation, interval estimation and prediction. Using the posterior distribution and the Bayes theorem, we make a single estimation about a value of interest, we can select a set of values with high probability, or we are able to make a estimation about future observations. In contrast to Bayesian inference, the solution methods in PPF analysis can only do at most two of these tasks.

As mentioned in the previous section, the MCS is the most widely used method for analyzing the PPF problem. However MCS does not consider the state variables as random variables with specific probability distributions. We are interested in taking into account this uncertainty over the state variables into the PPF analysis and one way to do this it is using Bayesian inference.

Bayesian inference uses a prior distribution and a likelihood function to obtain the

posterior probability distribution over the state variables, in contrast to MCS that employs repetitive deterministic solutions of the model and it then use a nonparametric density estimator for obtaining the probability distribution for the state variables. In this study, we show that the Bayesian inference methods (simulation-based Bayesian methods) can achieve proper results and it can take satisfactory computational times with respect to the state-of-the-art approaches.

Beyond power systems networks, Bayesian inference has also been used to study biology networks, for example in systems biology (Huelsenbeck et al., 2001; Wilkinson, 2007); and communication and transportation networks (Parry and Hazelton, 2013; Tebaldi and West, 1998), for analyzing network traffic counts, and for passenger flow assignment (Sun et al., 2015).

1.2 Aims

1.2.1 General aim

To develop a methodology for solving probabilistic power flow problems using Bayesian simulators.

1.2.2 Specific aims

1. To introduce a Bayesian inference framework for probabilistic power flow analysis, by specifying prior distributions over the state variables, a likelihood function that relates the state variables and the observations, and by providing posterior-distribution estimation based on modern sampling methods.

2. To develop a likelihood-free method for approximate Bayesian inference based on the power system information, for analyzing probabilistic power flow problems.

3. To validate the methods proposed for probabilistic power flow analysis and to compare their performance against state-of-the-art approaches using different test systems.

1.3 Contribution of this Research

The main contributions of this thesis include the following:

1. A Bayesian inference perspective for addressing the Probabilistic power flow problem is introduced.

2. We discuss alternatives for likelihood functions. Specifically, we present two likelihood functions: a multivariate Gaussian likelihood function and a multivariate t-Student likelihood function. From the PPF perspective, the first likelihood function means that all input random variables are modeled by Gaussian distributions. The last one can be seen as a probabilistic model from the observations that is tolerant to the negative effect of outliers. We also propose prior probability distributions over the state variables. Thus, the state variables are modeled as random variables and we also exploit an additional source of information that has not been used in simulation methods before. This contribution will be described in chapters 3 and 4.

3. We present a hierarchical Bayesian Modeling for the probabilistic power flow analysis. As commented before, we cannot use a simple multivariate Gaussian distribution for analyzing real power systems. Therefore, we propose to use a hierarchical Bayesian model to include the relationship between the output power of the wind turbine and the wind speed. The aim with this is to consider the wind speed as a random variable into the Bayesian PPF model. This hierarchical model is presented in chapter 3.

4. We use the most popular and powerful Bayesian frameworks for inferring the Bayesian probabilistic power flow problem, which are the MCMC and HMC methods. These methods are described in chapter 3.

5. We also provide a Bayesian methodology that incorporates the Jacobian computed from the power flow equations for enhancing the posterior distribution estimation of the state variables. We propose a new ABC method tailored to power systems, in which the Jacobian of the power flow equations of the PS is used to guide the search for more probable samples from the posterior distribution over the state variables. ABC methods applied to the PPF problems are detailed in chapter 4.

6. Finally, we present a methodology for choosing summary statistics in ABC methods. Specifically, we propose a new metric for comparing the distributions from the observed and simulated data in order to accept samples from the posterior distribution. This methodology is described in chapter 4.

1.4 Thesis Structure

This thesis is organized as follows. Chapter 2 gives a short introduction to PPF analysis and it also presents how to apply MCS to solve PPF problems. In chapter 3, we discuss a Bayesian modeling for PPF analysis and also present the two most popular Bayesian inference methods applied to this Bayesian PPF model. In chapter 4, we introduce likelihood-free Monte Carlo approaches or approximate Bayesian computation methods, as an alternative for solving PPF problems. Finally, in Chapter 5, we describe the conclusions and future works.

Chapter 2

Background

In this chapter, we give a short introduction to probabilistic power flow analysis. We expose the equations and variables of the PPF analysis. We also mention some methods proposed to solve the PPF problem. In addition, we present how to apply MCS to solve PPF problems. Finally, we briefly introduce the Bayesian inference and discuss some methodologies that have used Bayesian methods to take into account the uncertainty in PS.

2.1 Classic probabilistic power flow analysis

According to Su (2005) and given a network configuration, the power flow equations can be written as follows,

$$\mathbf{b} = \mathbf{g}(\mathbf{x}), \tag{2.1}$$

$$\mathbf{z} = \mathbf{h}(\mathbf{x}), \tag{2.2}$$

where $\mathbf{g}$ and $\mathbf{h}$ are nonlinear power flow equations. The vector $\mathbf{x} \in \mathbb{R}^{N_\mathbf{x}}$ includes the state variables, angles and voltages. If we assume a power system with N nodes and N_g nodes in PV nodes, the number of the unknown variables $\mathbf{x}$ is $N_\mathbf{x} = 2N - N_g - 2$, where the number of unknown angles and voltages would be equal to $N_\theta = N - 1$ and $N_V = N - N_g - 1$, respectively (Grainger and Stevenson, 1994). The vector $\mathbf{z}$ has elements given by the power flows through lines; and $\mathbf{b}$ is a vector with entries given by the net active and reactive powers injected, which are known. Besides, $\mathbf{b}$ depends on the powers generated (P_g) and loads (P_d). In PPF studies, P_g and P_d are modeled through probability distributions. It is common to use Gaussian, discrete and

Weibull distributions to model the uncertainty over the loads and powers generated (Soleimanpour and Mohammadi, 2013).

2.2 Related work

The first methods proposed to solve the PPF problem appeared in Borkowska (1974) and Dopazo et al. (1975) in the mid-1970s. Since then, several methods have been introduced with an interest on issues such as linearization, the effect of network outages, efficient and accurate methods for PPF, the extension of PPF analysis to three-phase systems in a distribution system and the inclusion of renewable energy into a PS or the modeling of uncertainty of the loads (Carmona-Delgado et al., 2015; Prusty and Jena, 2017).

From the studies of Borkowska (1974) and Dopazo et al. (1975), two philosophies have been accepted to analyze the uncertainty into the PS: stochastic power flow (SPF) (Dopazo et al., 1975; Vorsic et al., 1991) and probabilistic power flow (Allan et al., 1974; Allan and Al-Shakarchi, 1976, 1977; Borkowska, 1974). On one hand, SPF uses Eq. (2.1) and adds a white Gaussian noise $\boldsymbol{\xi}$ to model the uncertainty in the power flow equations, that is, $\mathbf{b} = \mathbf{g}(\mathbf{x}) + \boldsymbol{\xi}$. However, the SPF study is based on a least square estimation procedure applied to a linear approximation of the Eqs (2.1) and (2.2). This study does not define a likelihood function when renewable energy is considered into the PS. It also does not use prior distributions over the state variables to compute the probability distribution or the expected value and the variance of the solution of the probabilistic problem. On the other hand, the PPF approaches use hypothesis or mathematical assumptions of the power flow equations (see Eqs. (2.1) and (2.2)) to obtain efficient and accurate solutions of the probability distribution or statistical moments of the state variables and power flows.

In contrast to Eqs. (2.1) and (2.2), Aien et al. (2014) proposed to model the PPF problem assuming that $\mathbf{y} = \mathbf{f}(\mathbf{x})$, where $\mathbf{x}$ is the input variable vector that contains loads, network configuration and powers generated by distributed generation; $\mathbf{y}$ is the output vector that includes voltage magnitudes, angles and some powers generated; and $\mathbf{f}(\mathbf{x})$ determines the state of the system as a function of the input variables.

In Zhu (2015), the authors used the model of Eq. (2.1), and assumed that $\mathbf{b}$ is the

vector of input random variables and $\mathbf{x}$ is the vector of state variables; $\mathbf{b}$ is modeled by a multivariate Gaussian distribution. The mean vector and the covariance matrix for $\mathbf{b}$ are assumed known. Based on these assumptions, the mean vector for $\mathbf{x}$ can be computed through DPF methods and its covariance matrix, $\mathbf{\Sigma_x}$, can be obtained as $\mathbf{\Sigma_x} = \text{diag}\left(\mathbf{J}^\top \mathbf{\Sigma_b}^{-1} \mathbf{J}\right)$, where $\mathbf{J}$ is the Jacobian of the power flow equations; and $\Sigma_\mathbf{b}$ is the covariance matrix of the power injected, $\mathbf{b}$. According to the authors, $\mathbf{x}$ can be assumed to follow a multivariate Gaussian distribution. Such assumption is valid if all random variables are Gaussian distributed and the power flow equations are linearized around an operation point $\mathbf{x}_0$ (Dopazo et al., 1975). However, if the PS includes renewable energy, for example wind or solar energy, the random variables associated to these energies are non-Gaussian anymore (Aien et al., 2014), and the assumption about $\mathbf{x}$ is not necessarily Gaussian.

By using Bayesian inference, we are able to solve either the SPF or PPF problem. As mentioned before, we only need to define proper prior distributions for the state variables and a likelihood function for computing the posterior distribution over the state variables. In this project, we use the term "Bayesian PPF problem" to refer to the PPF problem when the state variables are considered as random variables and a full likelihood function is available. We call full likelihood function to indicate that the integration of the distributed generation is considered in the Bayesian PPF analysis

The solution methods for the PPF problem can be classified into three main categories: simulation-based methods, analytical methods or approximate methods (Carmona-Delgado et al., 2015; Gallego and Padilha-Feltrin, 2012; Le et al., 2013; Prusty and Jena, 2017). In the first category, the most widely used method is Monte Carlo simulation (Allan et al., 1981; Carmona-Delgado et al., 2015; Gallego and Padilha-Feltrin, 2012; Le et al., 2013; Vorsic et al., 1991), we can also find the Latin hypercube sampling (Yu et al., 2009) or uniform design sampling (Cai et al., 2014). Simulation methods are based on statistical sampling of input random variables and the propagation of these samples through repetitive deterministic solutions of the model. It obtains accurate and effective results, however, it is inefficient because the process requires a considerable number of deterministic solutions (Soleimanpour and Mohammadi, 2013; Zhang and Li, 2010). The second category includes approaches based on probabilistic analysis intervals (Briceno et al., 2012; Wang and Alvarado, 1992), multilinear variables (Allan and Leite da Silva, 1981; Leite da Silva and Arienti, 1990), convolution methods (Allan et al., 1981) or

cumulant methods (Fan et al., 2013; Le et al., 2013; Usaola, 2009, 2010; Zhang and Lee, 2004). Finally, the third category contains fuzzy theory (Kalesar and Seifi, 2010), point estimation methods (Morales and Perez-Ruiz, 2007; Saunders, 2013) or methods based on Taguchi's orthogonal arrays (Hong et al., 2016). The analytical and approximate methods are computationally more effective than methods based on simulation, however they require mathematical assumptions or approximations for feasible solutions (Gallego and Padilha-Feltrin, 2012; Soleimanpour and Mohammadi, 2013; Zhang and Li, 2010). Hence, the analytical and approximate methods may offer less accurate solutions than approaches based on simulation methods (Gallego and Padilha-Feltrin, 2012; Soleimanpour and Mohammadi, 2013; Zhang and Li, 2010).

2.2.1 Probabilistic power flow analysis based on Monte Carlo simulation

Classic probabilistic power flow analysis is based on repetitive solutions to the DLF problem (i.e. $\mathbf{g}(\mathbf{x}) - \mathbf{b} = \mathbf{0}$ from Eq. (2.1)). To solve the DLF problem, it is common to use the Newton Raphson method (Wood and Wollenberg, 1996). To obtain a solution, the DLF problem needs the specification of the real and reactive powers at PQ nodes; voltage magnitudes and real power injected at PV nodes and the network configuration. For a classic PPF problem, the generated power, demanded power, and available voltages are modeled as random variables with specific probability distributions. A large amount of samples are drawn by using the probability distributions for each of these random variables. Every one sample configuration is used to obtain a solution for the PPF problem. This process is repeated for the number of samples available. A nonparametric density estimator is used then to obtain the probability distribution of voltage magnitudes and angles at PQ buses; and angles at PV buses. This overall procedure is usually known as a Monte Carlo simulation (MCS).

2.3 Bayesian inference

The goal of Bayesian inference is to compute a posterior distribution over a random quantity $\mathbf{x}$

$$p(\mathbf{x}|\,\mathcal{D}) = \frac{p(\mathcal{D}|\,\mathbf{x})p(\mathbf{x})}{p(\mathcal{D})} = \frac{p(\mathcal{D}|\,\mathbf{x})p(\mathbf{x})}{\int p(\mathcal{D}|\,\mathbf{x})p(\mathbf{x})\,d\mathbf{x}}, \tag{2.3}$$

where $p(\mathbf{x})$ is the prior distribution (prior) for $\mathbf{x}$. The prior distribution is used to incorporate an initial hypothesis about $\mathbf{x}$. The term $p(\mathcal{D}|\mathbf{x})$ is the likelihood function, and it expresses how probable the observed data set is for different settings of $\mathbf{x}$. Notice that $p(\mathcal{D}|\mathbf{x})$ is not a probability distribution over $\mathbf{x}$ (Bishop, 2006; Murphy, 2012). The quantity $p(\mathbf{x}|\mathcal{D})$ is the posterior probability distribution (posterior) of the state variable $\mathbf{x}$ given observed data $\mathcal{D}$. The posterior quantifies the knowledge about the unknown variables and evaluate the uncertainty in $\mathbf{x}$ after observing $\mathcal{D}$ (Bishop, 2006; Murphy, 2012). In the Bayesian literature, there are several ways for computing the posterior distribution, which we can find approaches that use conjugate priors (Raiffa and Schlaifer, 2000, ch. 3), Laplace approximation (Murphy, 2012, pp. 255), variational inference (Bishop, 2006, ch. 10) and simulation-based Bayesian methods or Monte Carlo techniques (Bishop, 2006, ch. 11). In this project, we refer to the simulation-based Bayesian methods as Bayesian simulators.

2.3.1 Related work

As mentioned in the previous chapter, several studies have used Bayesian methods to take into consideration the uncertainty in PS. Specifically, these methodologies focus on modeling the uncertainty of the loads in a PPF problem (Carmona-Delgado et al., 2015; Mori and Jiang, 2009; Valverde et al., 2012). The authors in Carmona-Delgado et al. (2015); Mori and Jiang (2009); Valverde et al. (2012) employ a Gaussian mixture model to approximate non-Gaussian distributions of demands in a PS, where they also use prior distributions over the parameter of each mixture component. These parameters are computed by using the Expectation-Maximization algorithm (Bishop, 2006, pp. 438). Similarly in Dong et al. (2014), a Gaussian mixture model was used to approximate non-Gaussian distributions of loads. In contrast to Carmona-Delgado et al. (2015); Mori and Jiang (2009); Valverde et al. (2012), the authors employ variational Bayesian inference (Bishop, 2006, pp. 462) to compute the parameters of the mixture model. However, these methodologies do not employ Bayesian inference for obtaining the uncertainty over the state variables, since these studies only model the loads using Bayesian methods, and they then use MCS (repetitive deterministic solutions) to obtain the uncertainty over the state variables.

In the next chapter, we will present how the PPF problem can be seen as a Bayesian model. We will also briefly review how the likelihood-based Bayesian simulators can be applied to obtain samples from the posterior distribution of interest.

Chapter 3

Likelihood-based methods for Probabilistic Power Flow Analysis

In this chapter, we discuss a Bayesian model for the PPF analysis and also present the two most popular Bayesian inference methods applied to this Bayesian PPF model. Additionally, we also define prior distributions over the state variables and we address the feasibility of formulating a likelihood function of the power injected given the state variables. Finally, we describe the experiments and discuss the results obtained by applying the likelihood-based Bayesian simulators over two test systems: IEEE $\{6, 39\}$ bus test systems.

3.1 Bayesian Modeling for the Probabilistic power flow analysis

As commented in chapter 2, linear Gaussian models have been used to introduce the uncertainty in the power flow equations. Specifically, the study presented in Dopazo et al. (1975) is based on least square estimation procedure applied to a linear approximation of the Eqs (2.1) and (2.2). However, the authors do not consider the state variables as random variables. With the aim to contemplate the state variables as random variables, we will initially suppose to have the following linear power flow equation, which is a Taylor's series expansion of Eq (2.1),

$$\mathbf{b} = \mathbf{J}\mathbf{x} + \boldsymbol{\varepsilon}, \tag{3.1}$$

where $\mathbf{J}$ is the Jacobian of the power flow equations; and $\boldsymbol{\varepsilon}$ is a noise that models the uncertainty in $\mathbf{b}$. Now, we consider that the state variables are random variables and they are modeled by a prior distribution, which is given as,

$$\mathbf{x} \sim \mathcal{N}\left(\boldsymbol{\mu}_0, \boldsymbol{\Lambda}_0^{-1}\right), \tag{3.2}$$

where $\boldsymbol{\mu}_0$ and $\boldsymbol{\Lambda}_0$ are hyper-parameters that incorporate an initial knowledge about $\mathbf{x}$. These hyper-parameters can be assumed or estimated, since we could also put prior distributions on these hyper-parameters. Let us assume that $\boldsymbol{\varepsilon}$ is modeled by a Gaussian distribution, that is, $p(\boldsymbol{\varepsilon}) = \mathcal{N}(\mathbf{0}, \mathbf{L}^{-1})$, therefore, we have the following likelihood function,

$$p(\mathbf{b}|\,\mathbf{x}) = \mathcal{N}\left(\mathbf{b}|\,\mathbf{J}\mathbf{x}, \mathbf{L}^{-1}\right). \tag{3.3}$$

From Eqs. (3.1), (3.2) and (3.3), we can analytically obtain the posterior for $\mathbf{x}$ given $\mathbf{b}$ (for more details, see (Bishop, 2006, pp. 85)), as follows

$$p(\mathbf{x}|\,\mathbf{b}) = \mathcal{N}\left(\mathbf{x}|\,\boldsymbol{\Sigma}_{\mathbf{x}}\left(\boldsymbol{\Lambda}_0\boldsymbol{\mu}_0 + \mathbf{J}^\top\mathbf{L}\mathbf{b}\right), \boldsymbol{\Sigma}_{\mathbf{x}}\right), \tag{3.4}$$

where

$$\boldsymbol{\Sigma}_{\mathbf{x}} = \left(\boldsymbol{\Lambda}_0 + \mathbf{J}^\top\mathbf{L}\mathbf{J}\right)^{-1}.$$

As we can see in Eq. (3.4), the posterior $p(\mathbf{x}|\,\mathbf{b})$ uses the initial hypothesis over the state variable (see the product $\boldsymbol{\Lambda}_0\boldsymbol{\mu}_0$) and the information about the system (see the product $\mathbf{J}^\top\mathbf{L}\mathbf{b}$) to quantify the knowledge about the unknown state variables after observing $\mathbf{b}$. We also note, from Eq. (3.4), that the state variables are modeled by a multivariate Gaussian distribution, since the relationship between $\mathbf{b}$ and $\mathbf{x}$ is linear.

In real power systems, the relationship between $\mathbf{b}$ and $\mathbf{x}$ is non-linear, hence the posterior shown in Eq. (3.4) is not valid in real scenarios. Similarly, if we consider the integration of the distributed generation, for example wind or solar energy, the random variables associated to this energy are non-Gaussian anymore (Aien et al., 2014), and the assumption about $\mathbf{x}$ is not necessarily Gaussian.

Non-Gaussian distributions are employed to consider the uncertainty in loads or

powers generated; for example: loads have been modeled by multinomial or uniform distributions (Su, 2005), and renewable energy depends on a Weibull (Wind energy) or a beta (Solar energy) distribution (Aien et al., 2014).

The vector $\mathbf{b}$, in the Eq. (2.1), includes powers injected that depend on powers generated (P_g) and loads (P_d). When P_g and P_d are normally distributed, we simply use,

$$P_{g,d} \sim \mathcal{N}\left(\mu_{g,d}, \sigma_{g,p}^2\right), \tag{3.5}$$

where $P_{g,d}$ can be either a power generated or a load; $\mu_{g,d}$ is the the mean value and is assumed to be equal to the base load or the base power generated; and $\sigma_{g,p}$ is the standard deviation, which is set as the 5% of the mean value (Aien et al., 2014). [1] Therefore, $P_{g,d}$ is sampled from the distribution $\mathcal{N}\left(\mu_{g,d}, \sigma_{g,p}^2\right)$, in other words, "$z \sim$" is an abbreviation of "$p(z) =$" where $p(z)$ is the probability distribution for z.

On the other hand, the vector $\mathbf{b}$ can also be composed by loads modeled by discrete distributions, for example (Su, 2005)

$$p\left(P_{d_l}\right) = \mathrm{Dis}\left(\boldsymbol{\mu}_d, \boldsymbol{\pi}\right) = \begin{cases} \pi_1 & if \quad P_{d_l} = \mu_{d_1}, \\ \vdots & \vdots \quad\quad \vdots \\ \pi_m & if \quad P_{d_l} = \mu_{d_m}. \end{cases}, \tag{3.6}$$

where $p\left(P_{d_l}\right)$ is the probability distribution of the load μ_{d_l} at node l. π_j is the probability that the load Pd_j is equal to μ_{d_j}. We can rewrite Eq. (3.6) as follows,

$$p\left(P_{d_l}\right) = \pi_1 \delta\left(g_k\left(\mathbf{x}\right) + \mu_{d_1}\right) + \cdots + \pi_m \delta\left(g_k\left(\mathbf{x}\right) + \mu_{d_m}\right), \tag{3.7}$$

where π_j is the probability of each component, which is known. The values for π_j must satisfy $\sum_j \pi_j = 1$, and $\delta(\cdot)$ is the Dirac delta function. The presence of the Dirac delta mass function makes the application of many popular Bayesian techniques, such as Hamiltonian Monte Carlo in the model of Eq. (3.7) difficult (Mohamed et al., 2012). We proceed to change the Dirac delta mass functions by narrow Gaussian distributions

[1] This value can also be considered as a load or power generated forecasting error. Several approaches of load and energy demand forecasting have reported or notified a load forecasting error of about 5% (Wu et al., 2007).

(low variance), that is,

$$p\left(P_{d_l}\right) = \pi_1 \mathcal{N}\left(\mu_{d_1}, \sigma_q^2\right) + \cdots + \pi_m \mathcal{N}\left(\mu_{d_m}, \sigma_q^2\right), \tag{3.8}$$

where $\sigma_q^2 \to 0$. We use $\sigma_q^2 = 10^{-3}$ in this study. [2]

According to Soleimanpour and Mohammadi (2013), the uniform distribution is also used to model the loads in the PPF problem. The loads uniformly distributed can be modeled by,

$$P_{d_l} \sim \mathcal{U}\left(a_l, b_l\right), \tag{3.9}$$

where $\mathcal{U}$ is the uniform distribution for a continuous random variable P_{d_l} defined over a finite interval $P_{d_l} \in [a_l, b_l]$ where $b_l > a_l$.

On the other hand and in order to model the renewable energy penetration into the power system, we must define how is the relationship between the renewable output power and the renewable energy source. Let us assume we have a wind farm, where the output power from one wind turbine, P_w, in terms of the wind speed v_w, is given by

$$P_w\left(v_w\right) = \begin{cases} 0 & v_w \leq v_{cin}, \\ 0.5\rho A_w C_p v_w^3 & v_{cin} < v_w \leq v_r, \\ P_r & v_{cout} < v_w, \end{cases} \tag{3.10}$$

where v_{cin} is the cut-in wind speed; ρ is the air density; $A_w = \pi R^2$ is the area of the wind turbine rotor; R is the radius of the rotor; C_p is a coefficient of power, at which the wind turbine generator starts generating power (Soleimanpour and Mohammadi, 2013); v_r is the nominal rotational speed; P_r is the nominal wind power; and v_{cout} is the cut-out wind speed, at which the wind turbine generator is shut down for safety reasons (Soleimanpour and Mohammadi, 2013). Furthermore, v_w is the wind speed that is assumed to follow a Weibull distribution (Soleimanpour and Mohammadi, 2013),

$$p\left(v_w \left| a_{v_w}, b_{v_w}\right.\right) = \frac{b_{v_w}}{a_{v_w}} \left(\frac{v_w}{a_{v_w}}\right)^{b_{v_w}-1} e^{-\left(\frac{v_w}{a_{v_w}}\right)^{b_{v_w}}}, \tag{3.11}$$

[2] The selection of this value is discussed in section 3.3

where a_{v_w} and b_{v_w} are the scale and shape parameters, respectively. Another example of renewable energy is the solar cell generation, which is based on the solar radiation that varies in function of several factors such as environmental conditions, the time of day or the season (Aien et al., 2014). The solar radiation can be modeled by a beta distribution and the output power is deterministically related with the solar radiation through a similar model to the one presented in Eq. (3.10), see Aien et al. (2014) for more details.

Hierarchical Bayesian Modeling for the probabilistic power flow analysis

Having in mind the models exposed before, the PPF analysis can be introduced using a hierarchical Bayesian Model (HBM). We propose to use an HBM to include the relationship between the output power of renewable energies or distributed generations and the sources of these sustainable energies. An HBM is a multi-level model which consists of several levels of random variables or unknown variables (Murphy, 2012). Mathematically, an HBM is a simple modification to Eq. (2.3),

$$p\left(\mathbf{x},\mathbf{z}|\,\mathcal{D}\right)=\frac{p\left(\mathcal{D}|\,\mathbf{x},\mathbf{z}\right)p\left(\mathbf{x}|\,\mathbf{z}\right)p\left(\mathbf{z}\right)}{p\left(\mathcal{D}\right)}, \tag{3.12}$$

where $\mathbf{z}$ is a set of random variables or unknown parameters. The goal with the Eq. (3.12) is to integrate all the random variables with the observed data. We then apply Bayes theorem to obtain the posterior distribution over $\mathbf{x}$ and $\mathbf{z}$.

With an HBM, we are interested in modeling the dependence between random variables belonging to different levels as it happens in the case of distributed generation (see Eqs. (3.10) and (3.11)). A general structure of an HBM applied to PPF analysis can be seen in Fig. 3.1. From this figure, we present an HBM for PPF analysis as a directed graphical model,[3] where the vector $\mathbf{x}$ include the state variables; $\mathbf{b}$ is a vector that depends on P_g (powers generated), P_d (loads) and P_w (wind turbine output power); v_w is the wind speed, which can be modeled by a Weibull distribution; $\mathbf{f}_n$ is a simple adjustment to Eq. (2.1), that is, $\mathbf{f}_n = \mathbf{g}\left(\mathbf{x}_n\right) - \mathbf{b}_n$; $\mathcal{D}$ is an all-zeros vector (observed data); and N_s specifies the number of samples (simulations or observations) for each random variable. In Fig. 3.1, notice that the value $\mathbf{\Sigma}$ is not observed, so $\mathbf{\Sigma}$ can be

[3] A directed graphical model is an useful representation for probabilistic modeling (Murphy, 2012).

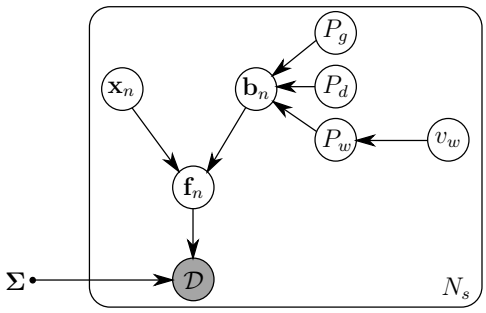


Figure 3.1: HBM for PPF analysis as a directed graphical model. The shaded circle is the observed data, and unshaded circles are random variables.

considered as hyper-parameter. Such variable is assumed known in this study.

From Fig. 3.1, a possible Bayesian PPF model is given as,

$$
\begin{aligned}
\mathcal{D}|\,\mathbf{f}_n, \boldsymbol{\Sigma} &\sim \mathcal{N}\left(\mathcal{D}|\,\mathbf{f}_n, \boldsymbol{\Sigma}\right), \\
\mathbf{f}_n &= \mathbf{g}\left(\mathbf{x}_n\right) - \mathbf{b}_n, \\
b_{n_k} &= P_{g_k} - P_{d_k} \quad \forall k, \\
P_{g_k} &\sim \begin{cases} \mathcal{N}\left(\mu_{g_k}, \sigma_{g_k}^2\right), \\ \text{Dis}\left(\boldsymbol{\mu}_g, \boldsymbol{\pi}_g\right), \end{cases} \\
P_{d_k} &\sim \begin{cases} \mathcal{N}\left(\mu_{d_k}, \sigma_{d_k}^2\right), \\ \mathcal{U}\left(a_l, b_l\right), \\ \text{Dis}\left(\boldsymbol{\mu}_d, \boldsymbol{\pi}_d\right), \\ P_w\left(v_w\right), \end{cases} \\
v_w &\sim p\left(v_w\,|a_{v_w}, b_{v_w}\right), \\
\mathbf{x}_n &\sim \mathcal{N}\left(\boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0\right),
\end{aligned}
\tag{3.13}
$$

where P_{g_k} can be modeled by a Gaussian or a discrete distribution (see Eqs (3.5) and (3.6), respectively); P_{d_k} can be a random variable modeled by a Gaussian, uniform, discrete distribution or it can be a wind farm output power that is often modeled as a negative load in the corresponding node (Aien et al., 2014); and $\mathcal{N}\left(\boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0\right)$ is the prior distribution over $\mathbf{x}$ with hyper-parameters $\boldsymbol{\mu}_0$ and $\boldsymbol{\Sigma}_0$. For PPF problems, $\boldsymbol{\mu}_0$ could be chosen as a previous DLF solution and the covariance matrix could be assumed known.

From Eq. (3.13) and Fig. 3.1, the likelihood function

$$p\left(\mathcal{D}|\,\mathbf{x},\mathbf{b}\right)=\prod_{n=1}^{N_s}\mathcal{N}\left(\mathcal{D}|\,\mathbf{f}_n,\boldsymbol{\Sigma}\right), \tag{3.14}$$

describes that the measurements are corrupted by an independent and identically distributed Gaussian noise with covariance matrix $\boldsymbol{\Sigma}$. Using (3.13) and (3.14) together with the Bayes' theorem, the posterior distribution for $\mathbf{x}$ and $\mathbf{b}$ has the form,

$$p\left(\mathbf{x},\mathbf{b}|\,\mathcal{D}\right)\propto\prod_{n=1}^{N_s}\mathcal{N}\left(\mathcal{D}|\,\mathbf{f}_n,\boldsymbol{\Sigma}\right)p\left(\mathbf{x}\right)p\left(\mathbf{b}\right), \tag{3.15}$$

where $\mathbf{b}$ includes the remaining random variables such as powers generated, loads and variables related to renewable energy and $p\left(\mathbf{b}\right)$ is the probability distribution that models the uncertainty for each random variable in $\mathbf{b}$.

In the next section, we explain how to obtain the posterior distribution over the state variables from the likelihood function and the prior distribution. For this, we present two likelihood-based Bayesian simulators: Markov Chain Monte Carlo and Hamiltonian Monte Carlo.

3.2 Likelihood-based Bayesian simulators

3.2.1 Markov Chain Monte Carlo

The most popular and powerful framework for likelihood-based Bayesian simulators is Markov Chain Monte Carlo (MCMC), which is used for sampling from high-dimensional distributions (Bishop, 2006; Murphy, 2012). The most general MCMC algorithm is the Metropolis Hastings algorithm. We use the term MCMC algorithm to refer to the Metropolis Hastings algorithm. The MCMC algorithm proposes a Markov chain on the state space $\mathbf{X}$, such that the stationary distribution ($p'\left(\mathbf{x}\right)$) for the Markov chain corresponds to the posterior distribution for the state variables. According to the Bayes theorem shown in Eq. (2.3), we are interested in sampling from

$$p\left(\mathbf{x}|\,\mathcal{D}\right)=p'\left(\mathbf{x}\right)=\frac{1}{Z_p}\tilde{p}\left(\mathbf{x}\right), \tag{3.16}$$

where $\tilde{p}\left(\mathbf{x}\right)$ is the product between the likelihood function and the prior distribution

(see Eq. (3.15)); and Z_p is the unknown normalizing constant. To sample from the posterior distribution, the MCMC algorithm proposes to move from the current state $\mathbf{x}$ to a new state $\mathbf{x}^*$ using a proposal distribution $q(\mathbf{x}^*|\,\mathbf{x})$, where the set of the state movements forms a Markov chain. The proposal distribution $q(\cdot)$ must be chosen so that we can easily evaluate it, and generate samples from it. Using q, we generate a candidate sample $\mathbf{x}^*$ and then accept the sample with probability,

$$\alpha = \min\left(1, \frac{p'(\mathbf{x}^*)\, q(\mathbf{x}|\,\mathbf{x}^*)}{p'(\mathbf{x})\, q(\mathbf{x}^*|\,\mathbf{x})}\right). \tag{3.17}$$

Using the expression in Eq. (3.16), then

$$\alpha = \min\left(1, \frac{\tilde{p}(\mathbf{x}^*)\, q(\mathbf{x}|\,\mathbf{x}^*)}{\tilde{p}(\mathbf{x})\, q(\mathbf{x}^*|\,\mathbf{x})}\right). \tag{3.18}$$

If the candidate sample is accepted from Eq. (3.18), then it becomes part of the proposed Markov chain, otherwise the candidate sample is rejected, and the current state is the sample previously accepted. The MCMC algorithm is summarized in algorithm 1.

Algorithm 1: MCMC algorithm

Initialize $\mathbf{x}_0$
for $i = 0, \ldots, N_s$ **do**
    Sample $\mathbf{x}^*$ from $q(\mathbf{x}^*|\,\mathbf{x}_i)$
    Compute α using the Eq. (3.18).
    Sample u from $\mathcal{U}(0, 1)$
    Set the sample to

$$\mathbf{x}_{i+1} = \begin{cases} \mathbf{x}^*, & u < \alpha \\ \mathbf{x}_i, & u \geq \alpha. \end{cases}$$

In the algorithm 1, N_s is the number of samples in the Markov chain; $\mathcal{U}(a, b)$ is an uniform distribution that depends on the parameters a and b; and u is a random number drawn from $\mathcal{U}(\cdot)$. For the MCMC, it is common to use a Gaussian distribution centered on the current state as proposal distribution. In this case, this algorithm is called random walk MCMC (Murphy, 2012).

3.2.2 Hamiltonian Monte Carlo

Another likelihood-based Bayesian simulator is Hamiltonian Monte Carlo (HMC), which uses the same philosophy of proposing a Markov chain on the state space $\mathbf{X}$ than MCMC. However HMC can efficiently avoid random walk behavior, that is, the motion of the accepted samples is random. HMC models the movement of a sample over the state space using a random initial momentum $\boldsymbol{\rho}$ and position $\mathbf{x}$ (Chen et al., 2014; Ding et al., 2014). HMC uses the energy and the dynamic of the system to determine the changes over time for $\boldsymbol{\rho}$ and $\mathbf{x}$. HMC introduces the Newton's laws of motion through the Hamiltonian function for formulating the motion of the samples deterministically, this function can be defined as (Chen et al., 2014),

$$H\left(\boldsymbol{\rho}, \mathbf{x}\right) = K\left(\boldsymbol{\rho}\right) + U\left(\mathbf{x}\right), \tag{3.19}$$

where the Hamiltonian $H\left(\boldsymbol{\rho}, \mathbf{x}\right)$ is a function of the kinetic energy $K\left(\boldsymbol{\rho}\right)$ and the potential energy $U\left(\mathbf{x}\right)$. Both energies, $K\left(\boldsymbol{\rho}\right)$ and $U\left(\mathbf{x}\right)$ are given as (Chen et al., 2014),

$$K\left(\boldsymbol{\rho}\right) = \frac{1}{2}\boldsymbol{\rho}^{\top}\mathbf{M}^{-1}\boldsymbol{\rho}, \tag{3.20}$$

$$U\left(\mathbf{x}\right) = -\log p\left(\mathbf{x}\right) - \sum_{d_i \in \mathcal{D}} \log p\left(d_i | \, \mathbf{x}\right), \tag{3.21}$$

where $\mathbf{M}$ is a mass matrix and according to Chen et al. (2014) it is often set as an identity matrix; and d_i is a component of $\mathcal{D}$. For the PPF problem, $\mathcal{D}$ can be computed by using the powers injected or the vector $\mathbf{b}$ in Eq. (2.1). From the above, the Hamiltonian dynamics can be expressed as

$$d\mathbf{x} = \boldsymbol{\rho} dt, \tag{3.22}$$

$$d\boldsymbol{\rho} = -\nabla U\left(\mathbf{x}\right). \tag{3.23}$$

For the HMC, the acceptance probability is computed using the above expressions, and it is given as

$$\alpha = \min\left(1, e^{H(\tilde{\boldsymbol{\rho}}, \tilde{\mathbf{x}}) - H(\boldsymbol{\rho}_i, \mathbf{x}_i)}\right), \tag{3.24}$$

where $\tilde{\boldsymbol{\rho}}$ and $\tilde{\mathbf{x}}$ are candidate samples for the momentum and the position, which are calculated using $K\left(\boldsymbol{\rho}\right)$ and $U\left(\mathbf{x}\right)$; $\boldsymbol{\rho}_i$ and $\mathbf{x}_i$ are the i-th momentum and i-th state variable that were accepted, respectively.

The HMC can be summarized as shown in algorithm 2.

Algorithm 2: HMC algorithm

Initialize $\mathbf{x}_0$
Set ε and m
for $i = 0, \ldots, N_s$ **do**
  Sample $\boldsymbol{\rho}_i \sim \mathcal{N}(\mathbf{0}, \mathbf{M})$
  $(\mathbf{x}_0, \boldsymbol{\rho}_0) = (\mathbf{x}_i, \boldsymbol{\rho}_i)$.
  $\boldsymbol{\rho}_0 \leftarrow \boldsymbol{\rho}_0 - \frac{\varepsilon}{2}\nabla U(\mathbf{x}_0)$
  for $j = 0, \ldots, m$ **do**
    $\mathbf{x}_j \leftarrow \mathbf{x}_{j-1} + \varepsilon \mathbf{M}^{-1}\boldsymbol{\rho}_{j-1}$
    $\boldsymbol{\rho}_j \leftarrow \boldsymbol{\rho}_{j-1} - \varepsilon \nabla U(\mathbf{x}_j)$
  $\boldsymbol{\rho}_m \leftarrow \boldsymbol{\rho}_m - \frac{\varepsilon}{2}\nabla U(\mathbf{x}_m)$
  $(\tilde{\mathbf{x}}, \tilde{\boldsymbol{\rho}}) = (\mathbf{x}_m, \boldsymbol{\rho}_m)$
  Sample $u \sim \mathcal{U}(0, 1)$
  $\alpha = \min\left(1, e^{H(\tilde{\boldsymbol{\rho}}, \tilde{\mathbf{x}}) - H(\boldsymbol{\rho}_i, \mathbf{x}_i)}\right)$
  if $u < \alpha$ **then**
    $\mathbf{x}_{i+1} = \tilde{\mathbf{x}}$
    $\boldsymbol{\rho}_{i+1} = \tilde{\boldsymbol{\rho}}$
  else
    $\mathbf{x}_{i+1} = \mathbf{x}_i$
    $\boldsymbol{\rho}_{i+1} = \boldsymbol{\rho}_i$

3.3 Results

In this section, we show four experiments that illustrate different aspects of this new approach for tackling the PPF problem. In the first experiment, we use the model of the load at bus six in the IEEE 6-bus test system presented in Su (2005) to estimate the value of σ_q^2 in the model in Eq. (3.8). In the second experiment, we presented the effects of prior distributions over the state variables on Bayesian PPF analysis. Specifically, we used two types of prior distributions for $\mathbf{x}$, which are: informative priors and non-informative priors. In the third experiment, we consider the PPF analysis over the IEEE 6-bus test system, which is described in appendix (A), and compare the performance between MCS, MCMC and the HMC. Finally, in the fourth experiment, we evaluate how the Bayesian simulators work when the number of nodes is increased. We use the IEEE 39-bus test system. For validation purposes, we compute relative errors with respect

to the values calculated using MCS, similar to Su (2005) and Morales and Perez-Ruiz (2007). The relative error is described in appendix B.

3.3.1 Selection of variance in narrow Gaussian distributions

As we commented before, we use narrow Gaussian distributions to approximate discrete distributions since the Dirac delta mass function makes difficult the application of HMC, as can be in the steps 6, 9 or 10 in algorithm 2. From Eq. (3.8), we present a narrow Gaussian mixture model with parameters $\left\{\boldsymbol{\mu}_d, \boldsymbol{\pi}, \sigma_q^2\right\}$. The two first parameters ($\boldsymbol{\mu}_d$ and $\boldsymbol{\pi}$) can be specified or assumed, see for example Su (2005) and Soleimanpour and Mohammadi (2013). The last one parameter must be estimated. Therefore, we have used the model of the load at bus six in the IEEE 6-bus test system shown in Su (2005) to generate the observed data, that is, from the discrete distribution of bus six and its parameters, we generate 1000 observations. We then drew 1000 samples from the our narrow Gaussian mixture model (see Eq. 3.8) with different values for $\sigma_q^2 = \{1, 0.1, 0.01, 0.001, 10^{-4}, 10^{-5}\}$. Finally, we computed a residual error between the observed and simulated data as can be seen in Fig. 3.2.

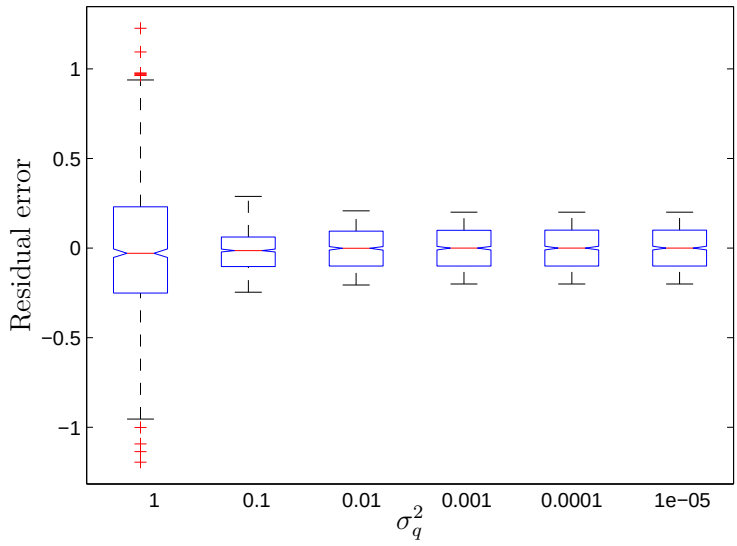


Figure 3.2: A boxplot for selecting the value for σ_q^2 in the mixture of narrow Gaussian distribution model in Eq. 3.8.

Fig. 3.2 shows that from $\sigma_q^2 = 10^{-3}$ the residual error between the observed and simulated data does not change. Therefore, we set $\sigma_q^2 = 10^{-3}$ for our following experiments.

3.3.2 Effect of Prior on Posterior on BPPF analysis

After having selected the variance in narrow Gaussian distributions to approximate discrete distributions, we are interested in showing the effects of prior distributions over the state variables on their posterior distributions on Bayesian PPF analysis. For this experiment, we used two types of prior distributions for $\mathbf{x}$, which are: informative priors and non-informative priors. An informative prior can express specific information about the variable. For example, from normal operating conditions of a PS, the magnitude of the voltage variables is close to one per unit, therefore, we can define a specific probability distribution for these variables. On the other hand, a non-informative prior provides vague or general information about the variable. For example, it is common that the angles of the voltage variables are initialized with zero values. In this experiment, we used the IEEE 6-bus test system and employed input Gaussian variables. We only modeled the angles as random variables and fixed the voltages in their deterministic solutions, since the input random variables are modeled by Gaussian distributions. We use Gaussian distributions as informative priors for each angle with mean equal to the DC power flow solution Zhu (2015), and standard deviation $\sigma_\theta = 0.1$. As non-informative priors, we employed Gaussian distributions with mean equal to zero, and standard deviation $\sigma_\theta = 0.1$. To observe the effects of prior distributions over the Bayesian PPF analysis, we first considered the angle at node 2 as a random variable and we fixed the four remaining angles in their deterministic solutions, we then applied HMC to obtain the posterior distribution of this variable. Next, we assumed that the angles at node 2 and 3 are random variables and fixed the three remaining angles in their deterministic solutions, and so on. For this Bayesian PPF problem, we employed the HMC approach with 20000 iterations and ignored the first 10000 samples. Finally, we computed a relative error (see Eq. (B.1)) between the deterministic solution and the posterior mean for each variable. We repeated the above procedure for 20 different input random variable configurations. Finally, we calculated the mean and one standard deviation of relative error. The Fig. 3.3 shows the relative errors for angles using the IEEE 6-bus test system when we considered informative and non-informative prior distributions.

From Fig. 3.3, we noted that in the first four cases, the results are close when we assumed informative priors and non-informative priors over the state variables. We also observed that the effect of the prior distributions on Bayesian PPF analysis can be seen when we considered five or more random variables. When we set non-informative priors over the state variables for these five angles, the relative error obtained using

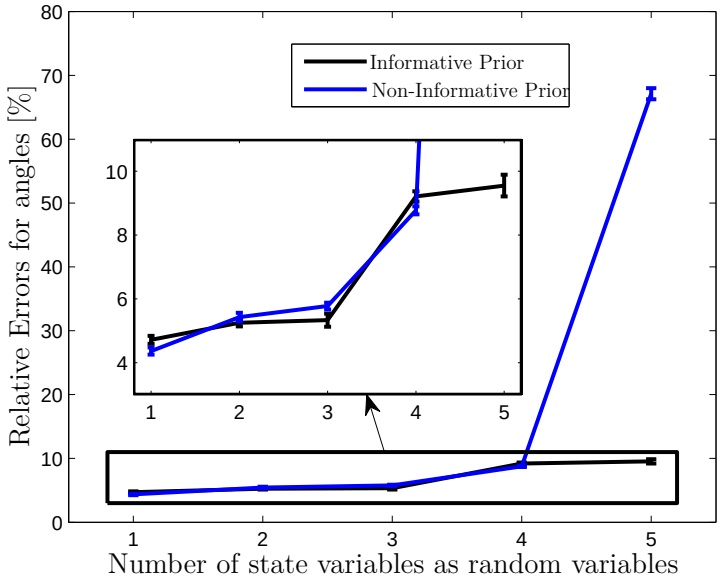


Figure 3.3: Relative errors for angles using the IEEE 6-bus test system when informative and non-informative prior distributions are considered.

non-informative priors is 67%. This value is inconsistent with the residual error value computed using informative priors, which is 9%. From the above, we recommended to use informative priors for analyzing Bayesian PPF problems. In the next experiments, we will employ informative priors over the state variables.

3.3.3 Results from IEEE 6-bus Test System

In this experiment, we compared the Bayesian inference methods against MCS for solving PPF problems, assuming a particular form for the likelihood. For this we used the IEEE 6-bus test system with constants and random variables employed in Su (2005). As example, we considered renewable energy in the system. Specifically, we add a wind farm at bus 4. Besides, we modeled the loads at nodes 5 and 6 as an uniform and discrete random variables, respectively. We set $P_{d_5} \sim \mathcal{U}(0.6, 0.8)$ and $p(P_{d_6}) = \text{Dis}([\ 0.7 \quad 0.6 \quad 0.5\ ], [\ 0.2 \quad 0.4 \quad 0.4\ ])$. P_{d_4} is a Gaussian random variable, similar to Su (2005). For the wind farm, we adopted the parameters used by Soleimanpour and Mohammadi (2013).[4] For the reactive loads at nodes 4, 5, 6 and power generated at nodes 2 and 3, we have also used Gaussian random variables as in Su (2005).

[4]For our experiments, we used $a = 15$ and $b = 2.5$. For the Eq. (3.10), we used $v_{cin} = 3$ m/s, $v_{cout} = 25$m/s, $v_r = 10.28$m/s, $C_p = 0.473$ and $R = 45$m.

From the above probability distributions, we generate 1000 samples for each random variable. [5]

Once we set the different random variables defined before, we used the Bayesian PPF model shown in Eq. (3.13) and the Bayesian framework described above to infer the posterior of $\mathbf{x} = \begin{bmatrix} \theta_2 & \theta_3 & \theta_4 & \theta_5 & \theta_6 & V_4 & V_5 & V_6 \end{bmatrix}^\top$ given $\mathcal{D}$, $p(\mathbf{x}|\mathcal{D})$. We use Gaussian distributions as prior distributions for $\mathbf{x}$, where its mean vector $\boldsymbol{\mu}_0$ was chosen as a previous DLF solution and the covariance matrix $\boldsymbol{\Sigma}_0$ was assumed to be a diagonal matrix with the following parameters: for the voltage variance, we use $\sigma_v^2 = 0.01$, and for the angles variance, we employ $\sigma_\theta^2 = 0.05$. We set $\boldsymbol{\Sigma} = 0.3\mathbf{I}$ in the likelihood function shown in Eq. (3.14). We then apply MCS, MCMC and HMC to infer the posterior of $\mathbf{x}$. We have used MATPOWER to implement the MCS. [6] MCMC was implemented on Matlab. HMC was implemented in stan on R. [7] For the MCMC and HMC, we employed 2000 iterations and ignored the first 1000 samples. For MCMC, we employed a multivariate Gaussian distribution as proposal distribution, where its covariance matrix is assumed to be a diagonal matrix with parameters, $\sigma_{v_q}^2 = 10^{-5}$ and $\sigma_{\theta_q}^2 = 10^{-4}$. For HMC, we use a step size $\varepsilon = 0.08$.

Fig. 3.4 shows the probability distribution of the state variables generated by the MCS, MCMC and HMC. The dashed red lines are the responses obtained by MCS. The solid magenta and black lines are the probability densities calculated by MCMC and HMC.

From Fig. 3.4, we note that MCMC poorly estimated the posteriors for each variable with respect to the probability distributions obtained by MCS. On the other hand, HMC correctly infers the posterior for θ_2, θ_3, θ_6, V_5 and V_6. For θ_4 and θ_5, HMC obtains samples around the mean computed by MCS, but the shape of the posterior for these two variables are not similar to the probability distributions obtained by MCS. Also, we note that HMC is not able to correctly estimate the posterior of variable V_4.

We have recorded the computation time (CT) required for each method to solve the PPF problems shown in Fig. 3.4. All simulations were conducted on an Intel Core i7

[5] We have used a number of samples of 1000 for our experiments with the goal of ensuring a coefficient of variation of 2% (for more details see Su (2005)). The coefficient of variation is defined as the ratio between the standard deviation and mean value.

[6] It is available at `http://www.pserc.cornell.edu/matpower/`

[7] Stan is a C++ library for Bayesian inference and it is available at `http://mc-stan.org/users/interfaces/rstan`

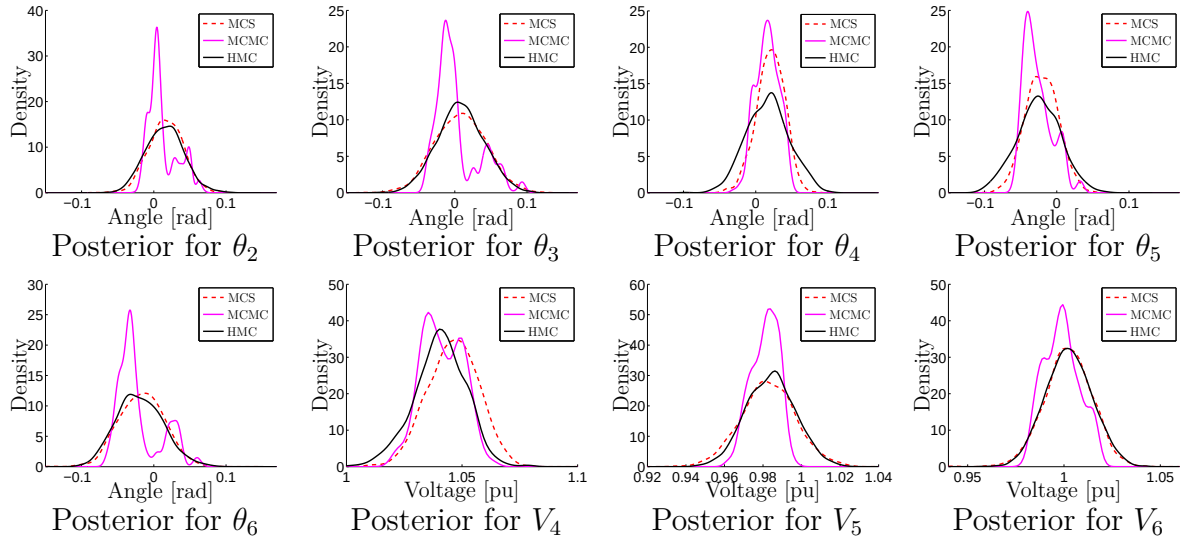


Figure 3.4: Posterior for each state variable. The dashed red lines are the responses obtained by MCS. The solid magenta and black lines are the probability densities calculated by MCMC and HMC. *Top row:* Posterior for θ_2, θ_3, θ_4 and θ_5. *Bottom row:* Posterior for θ_6, V_4, V_5 and V_6.

PC with a 2.1GHz processor. Table 3.1 lists the computation time took for the different methods. Notice that the computation times for MCMC and HMC are lower than the computation time took by MCS.

	MCS	MCMC	HMC
CT [s]	28.017	2.9328	0.9204

Table 3.1: Computation time (CT), in seconds [s], required by MCS, MCMC and HMC to solve the PPF problem over the IEEE 6-bus test system.

Next, we proceed to evaluate how close the mean and standard deviation of angles and voltages obtained by the Bayesian simulators are to the values computed by MCS, that is, we used the mean and standard deviation obtained by MCS as reference values. Using these values, we computed the relative error (see Eqs. (B.2) and (B.3)) for the MCMC and HMC. We generated 25 different sets of random variables, with 1000 samples for each variable. We then applied the MCMC and HMC to infer the posterior mean and standard deviation of the state variables using each random variable configuration. Finally, we computed 25 relative errors using the previous setup. Table 3.2 lists the mean and one standard deviation for the relative errors that compare the reference values obtained by MCS, and the estimated values using the Bayesian simulators based

on likelihood. As it would be expected, we observe that HMC gives better estimation results than MCMC. However, we note that the relative error for the standard deviation for the angles computed by MCMC is lower than the relative error obtained by HMC, this is due to HMC did not aptly estimate the shape of the distributions for θ_4 and θ_5.

IEEE 6		$\varepsilon_\theta[\%]$	$\varepsilon_V[\%]$
MCMC	μ	178.97 ± 140.34	0.4864 ± 0.2773
	σ	26.940 ± 13.651	26.351 ± 7.1361
HMC	μ	17.896 ± 11.797	0.2359 ± 0.0339
	σ	27.187 ± 2.8663	8.0704 ± 2.8766

Table 3.2: Relative error for voltages and angles using IEEE 6-bus test system.

3.3.4 Results from IEEE 39-bus Test System

Having shown the results for the IEEE 6-bus Test System, we proceed to evaluate the performance of our Bayesian approaches when the dimension of the unknown variables is increased. In this experiment, we use the IEEE 39-bus test system, which has 67 variables: 38 angles and 29 voltages. To analyze this system, we employ the random variables used in Soleimanpour and Mohammadi (2013), that is, we add a wind farm at bus 39, we also modeled the loads at nodes 18 and 26 as discrete random variables, and we also considered the loads at buses 4 and 29 as uniform random variables, but we do not contemplate correlated loads. The remaining variables are modeled by Gaussian distributions. From these random variables, we also generate 1000 samples for each variable. Similar to the previous experiment, we use a Gaussian prior distribution for $\mathbf{x}$, we use $\sigma_v^2 = 0.01$, and for the angles variance, we employ $\sigma_\theta^2 = 0.2$. We use $\mathbf{\Sigma} = \mathbf{I}$ in the likelihood function shown in Eq. (3.14). We then apply the MCS, MCMC and HMC. For MCMC and HMC, we used 2000 iterations and discarded the first 1000 samples. For MCMC, we utilized a multivariate Gaussian proposal distribution, with parameters, $\sigma_{v_q}^2 = 10^{-8}$ and $\sigma_{\theta_q}^2 = 10^{-5}$. For HMC, we use a step size $\varepsilon = 0.02$.

Fig. 3.5 displays the shape of the probability distributions for some variables using the IEEE 39-bus test system, since for space reasons we can not show the probability distribution of all the variables, however the results for the remaining variables are similar to those shown in this figure. For a comparison against MCS, we have randomly chosen the variables: θ_{12}, θ_{28}, θ_{24}, θ_{39}, V_1, V_4, V_{14} and V_{24}.

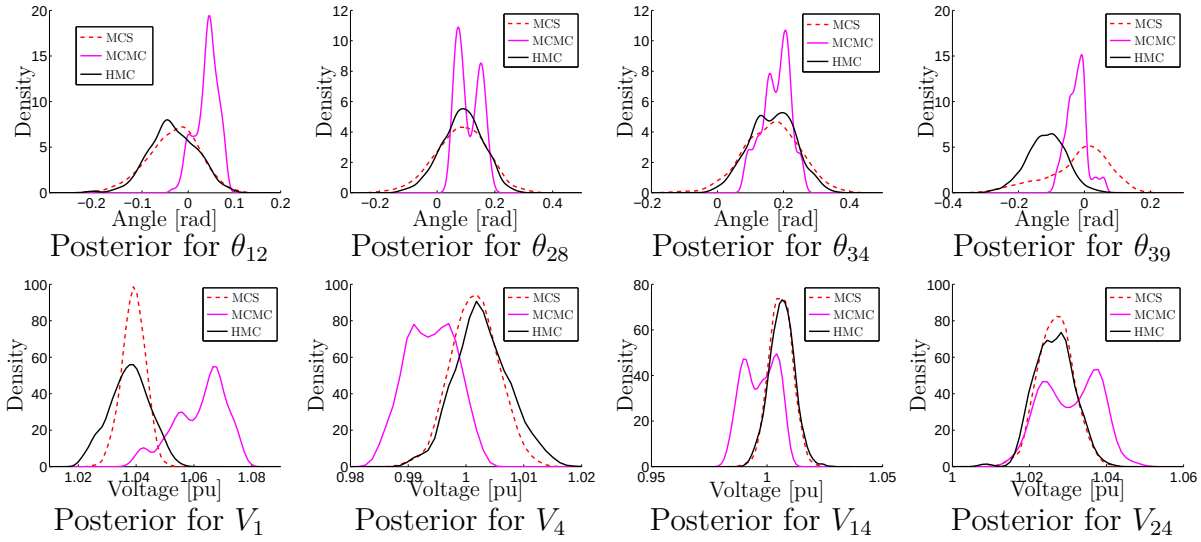


Figure 3.5: Posterior for some state variables. The dashed red line is the responses obtained by MCS. The solid magenta and black lines are the probability densities calculated by MCMC and HMC. *Top row:* Posterior for θ_{12}, θ_{28}, θ_{24} and θ_{39}. *Bottom row:* Posterior for V_1, V_4, V_{14} and V_{24}.

As can be seen in Fig. 3.5, the MCMC failed to infer satisfactory the posteriors for these variables. In contrast to HMC, the probability distributions for voltages are similar to those obtained by MCS. However, for some angles, HMC failed to correctly infer the probability distributions for these variables, as shown in the probability distributions for θ_{39} in Fig. 3.5.

We also recorded the computation time required for each method to solve the PPF problems shown in Fig. 3.5. Table 3.3 lists the computation time required for the MCS, MCMC and HMC to solve the PPF problem over the IEEE 39-bus test system. Notice that the computation times for MCMC and HMC are lower than the computation time took by MCS.

	MCS	MCMC	HMC
CT [s]	29.343	5.5380	12.412

Table 3.3: Computation time (CT), in seconds [s], required by MCS, MCMC and HMC to solve the PPF problem over the IEEE 39-bus test system.

Once we presented the comparison of the probability distributions obtained by MCMC and HMC against MCS, we proceeded to observe how close the results obtained by the

likelihood-based Bayesian simulators are to the values computed by MCS. Similar to what we had done at the previous experiment, we generated 25 different sets of random variables, we then apply the MCS, MCMC and HMC to compute the posterior mean and standard deviation of $\mathbf{x}$. Finally, we computed 25 relative errors between the values obtained by the Bayesian simulators and the results computed by MCS. These relative errors for angles and voltages using the posterior mean and standard deviation are summarized in Table 3.4.

IEEE 39		$\varepsilon_\theta[\%]$	$\varepsilon_V[\%]$
MCMC	μ	109.19 ± 28.320	0.9004 ± 0.1140
	σ	51.063 ± 3.6045	112.26 ± 12.739
HMC	μ	45.924 ± 8.2615	0.0772 ± 0.0074
	σ	27.457 ± 3.3922	45.556 ± 1.3228

Table 3.4: Relative error for voltages and angles using IEEE 39-bus test system.

As indicated in Table 3.4, the posterior mean and standard deviation computed by HMC are closer to the results obtained by MCS than those calculated by MCMC. However, it is necessary to highlight that the relative errors obtained by HMC for the angles, 45% for μ and 27% for σ, are due to HMC failed to correctly estimate the probability distributions for some of these variables, as it happened to θ_{39} in Fig. 3.5.

3.3.5 Discussion

The main purpose of this work was to include the uncertainty of state variables into the PPF problem, since this is not taken into account in simulation methods before, like MCS. We have compared our results using the Bayesian PPF model inferred with the likelihood-based Bayesian simulators against the results computed by MCS, how it is usually done in the PPF literature. However, we believe that it would not be possible a fair comparison between our methods and MCS, due to the MCS does not model the state variables as random variables. Likewise, our results are similar to the results obtained by MCS and we demonstrated that it is possible to include information about the system using Bayesian inference. Another point to mention is that the likelihood-based Bayesian simulators took less computation time than MCS for solving the PPF problems, see tables 3.1 and 3.3.

According to the results shown in Figs. 3.5 and 3.4, we validated that the use of Bayesian inference through the model of Eq. (3.13) and the HMC algorithm, which includes information about the system using the jacobian matrix, provides better results in comparison with those obtained by MCMC. This is due to the HMC can work in complex problems where MCMC fails (Gelman et al., 2013). However, it is necessary to specify or indicate that the proposed model (see Eq. (3.13)) should be improved to obtain even better results. In this chapter, we presented a model where its random variables and the observations are independent and identically distributed.

In the experiments performed in this chapter, we show how to integrate the renewable energy and non-Gaussian variables into the PS using a Bayesian model. We have demonstrated that the probability distributions obtained by our model are consistent with the results computed by MCS. However and as discussed previously, it is necessary to improve the model in order to obtain more accurate results, for example: the probability distribution for V_4 (in the IEEE 6-bus test system) and θ_{39} (in the IEEE 39-bus test system) must be improved.

On the other hand, we have approximated the model of a discrete random variable using a Gaussian mixture model. Thereby making it possible the use of HMC for obtaining the posterior distributions of the state variables.

Finally, we have shown an alternative model of the PPF problem using Bayesian inference, however a possible disadvantage of this Bayesian PPF modeling is that for each system a new model must be specified. Specifically, if the number of the unknown variables increases, or the number of the node increases, the number of the equations in the model presented in expression (3.13) also increases.

We have made available the R code that uses HMC based on Stan to analyze the IEEE 6 bus test at the following url:`https://github.com/cardazuluaga/PPFstanDemo.git`

Chapter 4

Likelihood-free methods for Probabilistic Power Flow Analysis

In the previous chapter, we presented a Bayesian PPF model that uses likelihood-based Bayesian simulators. An important issue about the previous chapter is the need to define a specific likelihood function for each PS. Therefore, this Bayesian model inferred by likelihood-based methods can not be generalized for different types of systems. In this chapter, we introduce likelihood-free Monte Carlo approaches or approximate Bayesian computation methods, as an alternative for solving the PPF problem. In practical Bayesian models, exact inference may be intractable due to different reasons: the likelihood function is expensive or intractable; or the normalization constant in Bayesian modeling is also intractable. The likelihood-free Bayesian simulators are employed to infer posterior distributions without having to evaluate likelihood functions, since they define the likelihood using simulator outputs.

4.1 Likelihood-free Bayesian simulators

4.1.1 Approximate Bayesian Computation

Approximate Bayesian Computation (ABC) are likelihood-free Monte Carlo methods. ABC methods can be employed to infer posterior distributions without having to evaluate likelihood functions (Toni et al., 2009; Wilkinson, 2013).

Given a prior distribution $p(\mathbf{x})$, the goal in ABC is to approximate the posterior distribution using the following model,

$$p\left(\mathbf{x}|\,\mathcal{D}\right) \propto g\left(\mathcal{D}|\,\mathbf{x}\right) p\left(\mathbf{x}\right), \tag{4.1}$$

where $g\left(\cdot\right)$ is a function or simulator that depends on the model. These likelihood-free algorithms simulate data using different parameters drawn from $p\left(\mathbf{x}\right)$ and compare summary statistics of the simulated data ($s\left(\mathcal{D}'\right)$) with summary statistics of the observed data ($s\left(\mathcal{D}\right)$). For this comparison, it is necessary to define a tolerance threshold that determines the accuracy of the algorithm, and a distance measure between the summary statistics $\rho\left(s\left(\mathcal{D}\right), s\left(\mathcal{D}'\right)\right)$. If $\rho\left(s\left(\mathcal{D}\right), s\left(\mathcal{D}'\right)\right) \leq \epsilon$, we then accept the state vector $\mathbf{x}$ drew from $p\left(\mathbf{x}\right)$ otherwise, the state vector is rejected. From Eq. (4.1), the principle behind ABC is to approximate the likelihood function using simulator outputs. For the PPF problem, the simulator corresponds to expression $\mathbf{g}\left(\mathbf{x}\right)$ in Eq. (2.1).

The most basic ABC algorithm is based on rejection sampling, and it is given by (Wilkinson, 2013)

Algorithm 3: ABC rejection

1 Draw $\mathbf{x}$ from $p\left(\mathbf{x}\right)$
2 Simulate $\mathcal{D}'$ using $\mathbf{g}\left(\mathcal{D}|\,\mathbf{x}\right)$
3 Accept $\mathbf{x}$ if $\rho\left(s\left(\mathcal{D}\right), s\left(\mathcal{D}'\right)\right) \leq \epsilon$

In this project, the algorithm 3 will be referred as ABC. It is important to mention that the empirical distribution over accepted samples for $\mathbf{x}$ is an approximation to the true posterior distribution. The approximation can be expressed by,

$$p\left(\mathbf{x}|\,\rho\left(s\left(\mathcal{D}\right), s\left(\mathcal{D}'\right)\right) \leq \epsilon\right).$$

If $\epsilon = 0$, the samples that we draw will come from the true posterior distribution, however the algorithm would need to perform more simulations for accepting any sample (Wilkinson, 2013). Since choosing $\epsilon = 0$ would be prohibitively expensive, the value of ϵ that we choose will affect the quality of the approximation.

4.1.2 Kernel Embeddings as summary statistics for ABC

According to Wilkinson (2013), other issues that arise around the ABC algorithm are how to choose a suitable distance measure $\rho(\cdot,\cdot)$, and how to choose the summary statistics $s(\cdot)$. With respect to s, this statistics function in likelihood-free models is usually

determined by expert domain knowledge. It is difficult to define a methodology for choosing or constructing the summary statistics (Park et al., 2015). Therefore, there is a motivation to construct the statistics automatically from the data (Blum et al., 2012; Jiang et al., 2015; Park et al., 2015)

Several methods have been proposed for determining summary statistics. We can find methods based on regression models (Fearnhead and Prangle, 2012; Jiang et al., 2015), methods based on information criterion (Blum, 2010; Blum et al., 2013; Nunes and Balding, 2010) or nonparametric ABC methods using kernel embeddings. These nonparametric ABC methods outperform other methods, ABC or semi-automatic ABC (Fearnhead and Prangle, 2012), for more details see Park et al. (2015). In this project, we will focus on the methodology based on kernel embeddings for choosing summary statistics.

Let us assume that we have two random samples $X = \{x_i\}_{i=1}^{N_x}$, and $Y = \{y_j\}_{j=1}^{N_y}$. In ABC, one of those samples would correspond to the real data ($\mathcal{D}$), whereas the other one would correspond to simulated data ($\mathcal{D}'$) from the PPF model, or the powers injected. As mentioned before, we need a way to decide if accepting the simulated data. A key idea introduced by Park et al. (2015) was to assume that the random samples X, and Y are drawn from probability measures P, and Q, respectively, and instead of comparing the distance between samples X, and Y, they propose to compare the distance between the probability measures P and Q.

The embedding of distributions in a Reproducing Kernel Hilbert Space (RKHS) is a methodology that allows us to compute distances between distributions by using kernel functions (Berlinet and Christine, 2004). According to Sriperumbudur et al. (2010), a metric $\gamma_k^2(P, Q)$ over the probability measures P and Q can be defined through a characteristic kernel[1] $k(x, x')$ as,

$$\gamma_k^2 (P, Q) = \left\| \int_M k (\cdot, x)\, dP (x) - \int_M k (\cdot, y)\, dQ (y) \right\|_{\mathcal{H}}^2 .$$

In this study, we use a Gaussian kernel as the characteristic kernel. If the distributions $P(x)$, and $Q(y)$ admit a density, then we have $dP(x) = p(x)dx$, and $dQ(y) = q(y)dy$,

[1] A characteristic kernel is a reproducing kernel for which $\gamma_k(P, Q) = 0 \iff P = Q, P, Q \in \mathcal{P}$, where $\mathcal{P}$ denotes the set of all Borel probability measures on a topological space $(M, \mathcal{A})$.

and an alternative expression for $\gamma_k^2(P,Q)$ can be written as

$$\begin{aligned}\gamma_k^2(P,Q) = &\int_M \int_M k(x,z)\, p(x)\, p(z)\, dxdz + \int_M \int_M k(z,y)\, q(z)\, q(y)\, dzdy \\ &- 2\int_M \int_M k(x,y)p(x)q(y)dxdy.\end{aligned} \tag{4.2}$$

In Park et al. (2015), the authors assume empirical distributions for P, and Q, this is $p(x) = \frac{1}{N_x}\sum_{i=1}^{N_x}\delta(x-x_i)$, and $q(y) = \frac{1}{N_y}\sum_{j=1}^{N_y}\delta(y-y_j)$. From the expressions for $p(x)$, and $q(y)$, the distance $\gamma_k^2(P,Q)$ is given by

$$\gamma_k^2(P,Q) = \frac{1}{N_x^2}\sum_{i,j=1}^{N_x} k(x_i,x_j) + \frac{1}{N_y^2}\sum_{i,j=1}^{N_y} k(y_i,y_j) - \frac{2}{N_xN_y}\sum_{i,j=1}^{N_x,N_y} k(x_i,y_j). \tag{4.3}$$

We refer to this distance as $\gamma_k^2(P,Q)_{\text{MMD}}$, since it is rooted in the Maximum Mean Discrepancy (MMD) concept developed in Gretton et al. (2007, 2012). After obtaining $\gamma_k^2(P,Q)_{\text{MMD}}$, Park et al. (2015) apply a second kernel that operates on probability measures, as follows

$$k_\epsilon(P_{\mathcal{D}}, Q_{\mathcal{D}'}) = \exp\left(-\frac{\gamma_k^2(P_{\mathcal{D}}, Q_{\mathcal{D}'})_{\text{MMD}}}{\epsilon}\right), \tag{4.4}$$

where ϵ is a positive parameter for the second kernel. $P_{\mathcal{D}}$ is the distribution associated to $\mathcal{D}$, and $Q_{\mathcal{D}'}$ is the distribution associated to $\mathcal{D}'$. In Park et al. (2015), the authors use an unbiased estimate for $\gamma_k^2(P,Q)_{\text{MMD}}$ in which the factors $1/N_x^2$, and $1/N_y^2$ are replaced for $1/(N_x(N_x-1))$, and $1/(N_y(N_y-1))$, respectively. Also, the innermost sum in each of the first two terms does not take into account the terms for which $i=j$.

Instead of assuming a discrete distribution for $p(x)$, and $q(y)$, in this study we propose to use a smooth estimate based on the Parzen-window density estimator for both densities. That is, we assume that $p(x)$, and $q(x)$ can be estimated using

$$\begin{aligned}\widehat{p}(x) &= \frac{1}{N_x}\sum_{m=1}^{N_x}\frac{1}{\left(2\pi h_p^2\right)^{D/2}}\exp\left(-\frac{\|x-x_m\|^2}{2h_p^2}\right), \\ \widehat{q}(y) &= \frac{1}{N_y}\sum_{n=1}^{N_y}\frac{1}{\left(2\pi h_q^2\right)^{D/2}}\exp\left(-\frac{\|y-y_n\|^2}{2h_q^2}\right),\end{aligned}$$

where h_p and h_q are the kernel bandwidths, and D is the dimensionality of the input space. If we use a Gaussian kernel with parameter Σ for $k(x,x')$, and the estimators

$\widehat{p}(x)$, and $\widehat{q}(y)$, a new distance between the distributions P and Q is easily obtained from expression (4.3) as follows

$$\gamma_k^2(P,Q) = \frac{1}{N_x^2}\sum_{i,j=1}^{N_x}\hat{k}(x_i,x_j;2\Sigma_p) + \frac{1}{N_y^2}\sum_{i,j=1}^{N_y}\hat{k}(y_i,y_j;2\Sigma_q) - \frac{2}{N_xN_y}\sum_{i,j=1}^{N_x,N_y}\hat{k}(x_i,y_j;\Sigma_p+\Sigma_q), \tag{4.5}$$

where

$$\hat{k}(x,x';S) = \frac{|\Sigma|^{1/2}}{|\Sigma+S|^{1/2}}\exp\left(-\frac{(x-x')^\top(\Sigma+S)^{-1}(x-x')}{2}\right).$$

In expression (4.5), $\Sigma_p = h_p^2\mathbf{I}$ and $\Sigma_q = h_q^2\mathbf{I}$. We refer to the metric in (4.5) as $\gamma_k^2(P,Q)_{\text{Parzen}}$.

As a distance measure $\gamma_k^2(P_{\mathcal{D}},Q_{\mathcal{D}'})$ we can use $\gamma_k^2(P_{\mathcal{D}},Q_{\mathcal{D}'})_{\text{MMD}}$ or $\gamma_k^2(P_{\mathcal{D}},Q_{\mathcal{D}'})_{\text{Parzen}}$. The algorithm proposed by Park et al. (2015) that employs kernel embeddings of probability measures in ABC using MMD is shown in Algorithm 4. If we use the metric $\gamma_k^2(P_{\mathcal{D}},Q_{\mathcal{D}'})_{\text{MMD}}$ on line 6 for the algorithm, we refer to the method as K2ABC. If we use the metric $\gamma_k^2(P_{\mathcal{D}},Q_{\mathcal{D}'})_{\text{Parzen}}$ instead, we refer to the method as PABC.

Algorithm 4: ABC based on Kernel Embeddings

1 **Input:** Observed data $\mathcal{D}$, prior distribution, threshold ϵ.
2 **Output:** Empirical posterior $\sum_{i=1}^{N_s}\tilde{w}_i\delta_{\mathbf{x}_i}$.
3 **for** $i = 1,\ldots,N_s$ **do**
4 Draw $\mathbf{x}_i$ from $p(\mathbf{x})$
5 Simulate $\mathcal{D}'$ using $\mathbf{g}(\cdot\,|\,\mathbf{x}_i)$
6 Compute $\widetilde{w}_i = \exp\left(-\frac{\gamma_k^2(P_{\mathcal{D}},Q_{\mathcal{D}'})}{\epsilon}\right)$
7 Normalize $\widetilde{w}_i$

From algorithm 4, in the step four, $\mathbf{x}_i$ corresponds to a new state variables, and $p(\mathbf{x})$ is the prior distribution over $\mathbf{x}$.

4.1.3 ABC MCMC

In algorithm 3, it is possible to have low acceptance rates when the prior distribution is not close to the posterior distribution. An alternative solution for this problem is

provided by a MCMC approach (Marjoram et al., 2003). ABC MCMC randomly explores the state space by modifying the current accepted samples. ABC MCMC proposes a new parameter value using a proposal distribution ($q(\cdot)$). This distribution must be chosen so that we can easily evaluate it, and generate samples from it. For more information about the proposal distribution and its parameters, see Murphy (2012, pp. 850). Algorithm 5 shows how the ABC MCMC method can be implemented.

Algorithm 5: ABC MCMC

1 Initialize $\mathbf{x}_0$
2 **for** $i = 1, \ldots, N_s$ **do**
3 Draw $\mathbf{x}^*$ from $q(\mathbf{x}|\,\mathbf{x}_i)$
4 Simulate $\mathcal{D}'$ using $\mathbf{g}(\mathcal{D}|\,\mathbf{x}^*)$
5 if $\rho(\mathcal{D}, \mathcal{D}') \leq \epsilon$
6 Accept $\mathbf{x}_{i+1} = \mathbf{x}^*$ with probability
7 $\alpha = \min\left(1, \frac{p(\mathbf{x}^*)q(\mathbf{x}_i|\mathbf{x}^*)}{p(\mathbf{x}_i)q(\mathbf{x}^*|\mathbf{x}_i)}\right)$
8 Otherwise $\mathbf{x}_{i+1} = \mathbf{x}_i$

Algorithm 5 also obtains samples from an approximated posterior over $\mathbf{x}$ (Toni et al., 2009). However, according to Toni et al. (2009), ABC MCMC may get stuck in regions of low probability for the accepted samples. For avoiding this, we propose to include information of the PS during the ABC inference stage.

4.1.4 Jacobian ABC for probabilistic power flow problems

To deal with the problems mentioned before, we propose to use the Jacobian of the power flow equations as part of the ABC MCMC method. We employ the Jacobian matrix information to do an improved state space exploration. From algorithm 5, notice that the acceptance probability (step 7) is given by

$$\alpha = \min\left(1, \frac{p(\mathbf{x}^*)\, q(\mathbf{x}_i|\,\mathbf{x}^*)}{p(\mathbf{x}_i)\, q(\mathbf{x}^*|\,\mathbf{x}_i)}\right), \tag{4.6}$$

where q is a proposal distribution or transition kernel, $\mathbf{x}^*$ is a candidate state variable drew from q, and $\mathbf{x}_i$ is the ith set of state variables accepted. If we assume a symmetric proposal distribution, that is, $q(\mathbf{x}^*|\,\mathbf{x}_i) = q(\mathbf{x}_i|\,\mathbf{x}^*)$, the acceptance probability is given

by the following expression,

$$\alpha = \min\left(1, \frac{p\left(\mathbf{x}^*\right)}{p\left(\mathbf{x}_i\right)}\right). \tag{4.7}$$

We note that if $\mathbf{x}^*$ is more probable than $\mathbf{x}_i$, we definitely accept the candidate state variable $\mathbf{x}^*$ (since $\frac{p(\mathbf{x}^*)}{p(\mathbf{x}_i)} > 1$). Due to $\mathbf{x}^*$ is drawn from q, we propose to use a multivariate Gaussian distribution as symmetric proposal distribution, where its mean value is updated as follows (Bracale et al., 2013; Meeds and Welling, 2015),

$$\mathbf{x}_{i+1} = \mathbf{x}^* + \mathbf{J}_i^{-1}\left(\mathcal{D} - \mathcal{D}'\right), \tag{4.8}$$

where the matrix $\mathbf{J}_i : \mathbb{R}^{N_\mathbf{x}} \to \mathbb{R}^{N_\mathbf{x}}$ is the Jacobian of $\mathbf{g}\left(\mathbf{x}\right)$, that is, $\mathbf{J}_i = \left.\frac{\partial \mathbf{g}(\mathbf{x})}{\partial \mathbf{x}}\right|_{\mathbf{x}=\mathbf{x}_i}$ (Grainger and Stevenson, 1994); and $\mathcal{D} - \mathcal{D}'$ is a vector of relative errors between observed and simulated data. Eq. (4.8) can be seen as a correction step, which seeks that $\mathcal{D}'$ equals $\mathcal{D}$ through $\mathbf{x}^*$. We also need to define the covariance matrix for the distribution q, $\mathbf{\Sigma}_\mathbf{x}$. In this paper, we use a diagonal matrix with N_θ elements $\sigma_{\theta_q}^2$ for angles and N_V diagonal entries $\sigma_{v_q}^2$ for voltages. The basic idea of using Eq. (4.8) is to move from a current state $\mathbf{x}_i$ to a more probable new state $\mathbf{x}^*$. Therefore, the steps six to eight of algorithm 5 are unnecessary. We refer to this new ABC algorithm as JABC. The JABC algorithm can be summarized as follows,

Algorithm 6: JABC

1 Draw $\mathbf{x}^*$ from $q\left(\mathbf{x}\middle|\,\mathbf{x}_i, \Sigma_\mathbf{x}\right)$
2 Simulate $\mathcal{D}'$ using $\mathbf{g}\left(\mathcal{D}\middle|\,\mathbf{x}^*\right)$
3 if $\rho\left(\mathcal{D}, \mathcal{D}'\right) \leq \epsilon$
4 $\mathbf{x}_{i+1} = \mathbf{x}^* + \mathbf{J}_i^{-1}\left(\mathcal{D} - \mathcal{D}'\right)$

For the algorithm 6, it is necessary to choose adequately the initial condition $\mathbf{x}_0$. In this study, we take an all-ones vector for voltages and the solution of DC power flow algorithm for angles.

4.1.5 Extension to ABC SMC

The disadvantages of ABC, such as: the selection of ϵ , the choice of summary statistics or accepted samples with low probability, can be avoided using an ABC algorithm based on sequential Monte Carlo methods (ABC SMC) proposed by Sisson et al. (2007). The goal of ABC SMC is to obtain an approximation of the true posterior using a series of sequential steps, expressed by $p\left(\theta\middle|\,\rho\left(\mathcal{D}, \mathcal{D}'\right) \leq \epsilon_t\right)$, for $i = 1, \cdots, T$, where ϵ_t is a threshold

that decreases in each step ($\epsilon_1, >, \ldots, \epsilon_t, >, \ldots, > \epsilon_T$), thus it refines the approximation toward the target distribution.

ABC SMC has a first stage based on ABC rejection. We can replace this stage with JABC, K2ABC or PABC, leading to what we call in this document as the JABC SMC, K2ABC SMC or PABC SMC, respectively. Details of the ABC SMC method can be found in Toni et al. (2009). A qualitative comparison (from the PPF analysis viewpoint) of the ABC algorithms is summarized in Table 4.1.

Algorithm	Distance-free	Summary Statistics	Threshold
ABC	✗	✓	✓
ABC MCMC	✗	✓	✓
K2ABC	✓	✗	✗
PABC	✓	✗	✗
JABC	✗	✗	✓

Table 4.1: Comparison of ABC algorithms.

4.2 Results

In this section, we present several numerical problems to illustrate different properties of the likelihood-free methods for dealing with PPF problems. In the first experiment, we show an example where the likelihood functions are actually known and compare the performance between MCS, Hamiltonian Monte Carlo, and ABC in an IEEE 6-bus test system. We then compare the performance of the ABC methods to analyze the IEEE 39 bus test system with non-Gaussian random variables (see the experiment presented in subsection 3.3.4). For the second experiment, we evaluate how the ABC algorithms work, including the sequential Monte Carlo variants, compare with MCS in terms of the number of nodes of a PS. We use the IEEE $\{6, 39, 118\}$-bus test systems. For the third experiment, we compare the ABC methods against other methodologies, we specifically use the point-estimation method based on two points (Su, 2005) and the Taguchi method (TM) (Hong et al., 2016), to analyze PPF problems. For the fourth experiment, we show the performance of the ABC methods when including renewable sources of energy in the IEEE 39-bus test system.

4.2.1 PPF analysis using MCS and Bayesian inference

With the aim of comparing the performance of Bayesian inference methods against to MCS for solving PPF problems, we have used the IEEE 6-bus test system, we have assumed a particular form for the likelihood function and we also assumed to know the true posterior of $\mathbf{x}$. To define such posterior, we use the input random variables as in Su (2005) (no uncertainty in line parameters) and employ MCS for obtaining the probability distribution of $\mathbf{x}$, that we use as the true posterior of $\mathbf{x}$. We initially drew 1000 samples from this posterior and computed $\{\mathbf{b}_i\}_{i=1}^{1000}$ through Eq. (2.1). [2] To calculate $\mathbf{b}_i$, which corresponds to $\mathcal{D}$ for the ABC methods, we used two different likelihood functions: a multivariate Gaussian likelihood function (MGLF) with mean $\mathbf{g}(\mathbf{x})$ and covariance matrix $\boldsymbol{\Sigma} = 3 \times 10^{-4}\mathbf{I}$, that is,

$$p(\mathbf{b}|\,\mathbf{x}) = \prod_{i=1}^{1000} \mathcal{N}(\mathbf{b}_i|\,\mathbf{g}(\mathbf{x}), \boldsymbol{\Sigma}),$$

where $\mathcal{N}(\mathbf{b}|\,\mathbf{g}(\mathbf{x}), \boldsymbol{\Sigma})$ can be seen as,

$$\mathcal{N}(\mathbf{b}|\,\mathbf{g}(\mathbf{x}), \boldsymbol{\Sigma}) = \frac{1}{(2\pi)^{N_{\mathbf{x}}/2}\,|\boldsymbol{\Sigma}|^{1/2}} \exp\left(-\frac{1}{2}(\mathbf{b}-\mathbf{g}(\mathbf{x}))^{\top}\boldsymbol{\Sigma}^{-1}(\mathbf{b}-\mathbf{g}(\mathbf{x}))\right).$$

On the other hand, we also used a multivariate t-Student likelihood function (MTLF) with mean $\mathbf{g}(\mathbf{x})$, precision matrix $\boldsymbol{\Lambda} = \boldsymbol{\Sigma}^{-1}$ and 5 degrees of freedom ($\nu = 5$), which is defined as,

$$p(\mathbf{b}|\,\mathbf{x}) = \prod_{i=1}^{1000} \mathcal{S}(\mathbf{b}_i|\,\mathbf{g}(\mathbf{x}), \boldsymbol{\Lambda}, \nu),$$

where $\mathcal{S}(\mathbf{b}|\,\mathbf{g}(\mathbf{x}), \boldsymbol{\Lambda}, \nu)$ is given by,

$$\mathcal{S}(\mathbf{b}|\,\mathbf{g}(\mathbf{x}), \boldsymbol{\Lambda}, \nu) = \frac{\Gamma(\nu/2 + N_{\mathbf{x}}/2)}{\Gamma(\nu/2)} \frac{|\boldsymbol{\Lambda}|^{1/2}}{(\nu\pi)^{N_{\mathbf{x}}/2}} \left[1 + \frac{\boldsymbol{\Delta}^2}{\nu}\right]^{-\nu/2-1/2},$$

where $\boldsymbol{\Delta}^2 = (\mathbf{b}-\mathbf{g}(\mathbf{x}))^{\top}\boldsymbol{\Lambda}(\mathbf{b}-\mathbf{g}(\mathbf{x}))$ and $\Gamma(x)$ is the gamma function defined by

[2] We have used a number of samples of 1000 for our experiments with the goal of ensuring a coefficient of variation of 2% (for more details see Su (2005)). The coefficient of variation is defined as the ratio between the standard deviation and mean value.

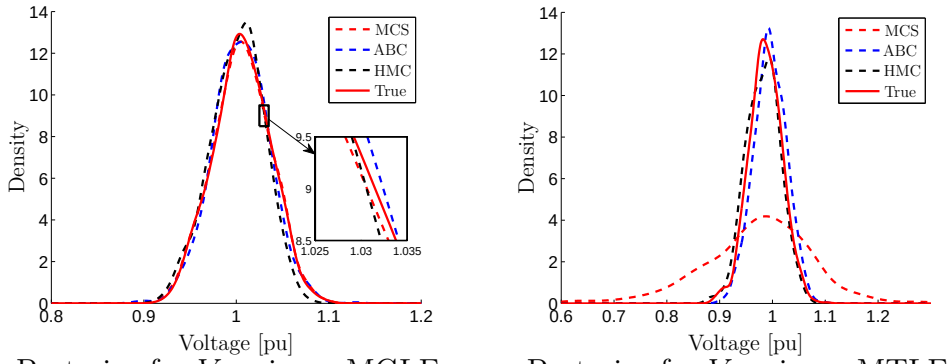
$\Gamma(x) \equiv \int_0^\infty u^{x-1}e^{-u}du$.

From the PPF perspective, the MGLF means that each $\mathbf{b}_i$ is modeled by Gaussian distributions (input Gaussian variables). If we assume that the relationship between $\mathbf{b}_i$ and $\mathbf{x}$ is linear, we can have a similar likelihood used by Dopazo et al. (1975). On the other hand, a MTLF can be considered as a probabilistic model from the observations that it is tolerant to the negative effect of outliers, for example: due to a measurement equipment malfunction of the input variables.

Once we generated the different datasets using these two likelihood functions, we used the Bayesian inference framework described above to make inference for the posterior $p(\mathbf{x}|\,\mathbf{b})$. For each voltage, we use Gaussian distributions as prior distributions with mean equal to one, and variance $\sigma_v^2 = 0.0015$. We use uniform distributions as prior distributions for each angle, i. e., $\theta_i \sim \mathcal{U}(a_i, b_i)$. Parameters a_i and b_i are computed as $a_i = \theta_i^{DC} - \Delta\theta$ and $b_i = \theta_i^{DC} + \Delta\theta$, where θ_i^{DC} is the ith DC power flow solution (Zhu, 2015); and $\Delta\theta$ quantifies the error present in the solution obtained by the DC power flow algorithm with respect to the AC power flow solution (Eq. (2.1)). We chose $\Delta\theta$ equal to 0.07. We employed ABC and MCS to infer the posterior of $\mathbf{x}$. For ABC, we put $\epsilon = 0.7$, the simulation function was set as $\mathbf{g}(\mathbf{x})$ and we ran one simulation by each $\mathbf{b}_i$. As summary statistics in ABC (see step 3 in algorithm 3), we compute: the losses in the PS, the total reactive power for loads, the average of the quadratic active and reactive power for loads. [3] We also compared these two methods against HMC. Fig 4.1 shows the probability distribution of V_6 generated by the MCS, ABC and HMC, when we assumed a MGLF (left) and a MTLF (right).

Fig. 4.1 (left) shows the posteriors inferred when a MGLF was assumed. We observe that MCS, ABC and HMC obtain satisfactory posteriors. When we used a MTLF (see Fig. 4.1 (right)), which is an intractable likelihood, we note that MCS did not infer appropriately the posterior for V_6, however, ABC and HMC properly estimated the posterior of V_6. From these results, we see that the MCS is affected by outliers, convergence properties and high variability. In contrast to HMC, ABC did not need a particular form for the likelihood to obtain the posterior of V_6. On the other hand, these results confirm

[3] These summary statistics can be computed as: $P_L = \sum_{i=1}^{N} P_i$, $Q_T = \sum_{i=1}^{N} Q_i$, $Q_T^2 = \frac{1}{N}\sum_{i=1}^{N} Q_i^2$ and $P_T^2 = \frac{1}{N}\sum_{i=1}^{N} P_i^2$. N is the number of the nodes; P_i and Q_i are the active and reactive powers injected and can be obtained by using the Eq. (2.1).



Posterior for V_6 using a MGLF Posterior for V_6 using a MTLF

Figure 4.1: Posterior for V_6 when two likelihood functions are considered. The dashed red, blue and black lines are the responses obtained by MCS, ABC and HMC. The solid red lines are the true posteriors. *Left:* Posterior for V_6 by using a MGLF. *Right:* Posterior for V_6 by using MTLF.

the importance of including additional source of information (a prior over $\mathbf{x}$) to analyze the uncertainty of $\mathbf{x}$, that is, both HMC and ABC updated the initial information about $\mathbf{x}$ using the likelihood functions to obtain the posteriors shown in Fig 4.1, whereas MCS does not. We observed the results for other variables of $\mathbf{x}$, and they are very similar to the results of V_6.

Next, we proceeded to use the MCS and the likelihood-free Bayesian simulators such as: ABC, K2ABC, PABC, ABCMCMC and ABC SMC to analyze the experiment shown in Fig. 3.5. For that experiment, which uses the IEEE 39-bus test system, we select and employ 10 summary statistics, four of them are the ones computed in the previous experiment, and the six remaining statistics are: the maximum and minimum values of the active and reactive powers injected; and the average of the active and reactive powers injected. [4] We generate the same datasets and also use the informative prior distributions for $\mathbf{x}$ as we specified in the subsection 3.3.4. For ABC and ABC MCMC, we set $\epsilon = 20$. For ABC SMC, we have employed a series of sequential steps $\{\epsilon_t\}_{t=1}^{T} = \{30.0, 28.0, 25.0, 22.0, 20.0\}$. For K2ABC and PABC, we use $\epsilon = 1.0$, this value was chosen using a sensitivity analysis for this parameter, similar to Park et al. (2015). We also use, for K2ABC and PABC, a kernel bandwidth optimization approach based on minimizing the mean integrated square error between the observed and simulated

[4] These six remaining statistics are: $\max_P \mathbf{b}$, $\min_P \mathbf{b}$, $\max_Q \mathbf{b}$, $\min_Q \mathbf{b}$, $\tilde{Q} = \frac{1}{N}\sum_{i=1}^{N} Q_i$ and $\tilde{P} = \frac{1}{N}\sum_{i=1}^{N} P_i$.

data, to define the values of Σ, h_x and h_y. For more details about this kernel bandwidth optimization approach, see Shimazaki and Shinomoto (2010). In Fig. 4.2, we present the posterior for θ_{12}, θ_{28}, θ_{24} and θ_{39} (see in the top row); and for V_1, V_4, V_{14} and V_{24} (see in the bottom row). From the Fig 4.2 the dashed red, blue magenta, red, gray and black lines are the probability densities obtained by MCS, ABC, K2ABC, PABC, ABC MCMC and ABC SMC, respectively.

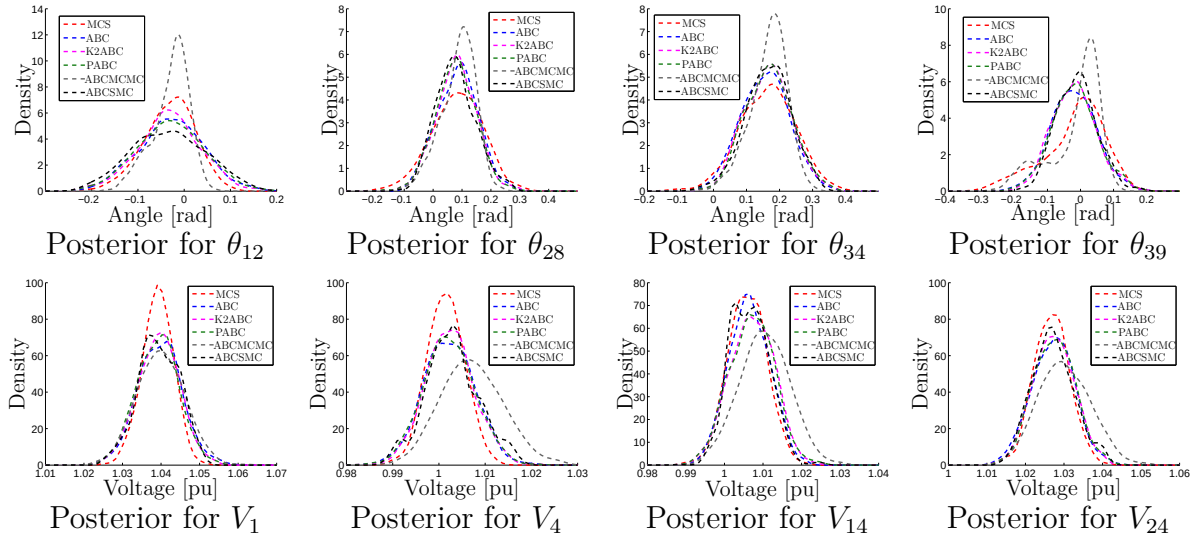


Figure 4.2: Posterior for some state variables. The dashed red line is the responses obtained by MCS. The dashed red, blue magenta, red, gray and black lines are the probability densities calculated by ABC, K2ABC, PABC, ABC MCMC and ABC SMC, respectively. *Top row:* Posterior for θ_{12}, θ_{28}, θ_{24} and θ_{39}. *Bottom row:* Posterior for V_1, V_4, V_{14} and V_{24}.

From Fig. 4.2, the results obtained by ABC, K2ABC, PABC and ABC SMC are similar. These results, for K2ABC and PABC, confirm the use of Kernel Embeddings as summary statistics, since for ABC we manually chose 10 summary statistics. From this figure, we also observed that the results computed by ABC MCMC are acceptable, in some cases the shape of the posterior distributions does not match to the one obtained by MCS.

To show how close the mean and standard deviation of angles and voltages obtained by the likelihood-free methods are to the values computed by MCS, we have repeated the experiment shown in Fig. 4.2 using 25 different sets of random variables. We then apply each method and computed 25 relative errors between the values obtained by the Bayesian simulators and the results computed by MCS. Table 4.2 shows the relative

errors for angles and voltages using the posterior mean and standard deviation.

IEEE 39		$\varepsilon_\theta[\%]$	$\varepsilon_V[\%]$
ABC	μ	7.4114 ± 2.2498	$\mathbf{0.0585 \pm 0.0040}$
	σ	16.859 ± 1.6281	49.251 ± 1.8405
K2ABC	μ	7.5240 ± 1.8387	0.0602 ± 0.0039
	σ	17.177 ± 0.7613	$\mathbf{49.120 \pm 1.9359}$
PABC	μ	$\mathbf{7.3427 \pm 1.8813}$	0.0608 ± 0.0037
	σ	17.215 ± 0.7060	49.178 ± 1.8440
ABC MCMC	μ	12.457 ± 2.3025	0.3609 ± 0.0375
	σ	26.310 ± 1.0709	68.138 ± 2.2167
ABC SMC	μ	21.515 ± 7.8962	0.0751 ± 0.0106
	σ	$\mathbf{16.108 \pm 1.6962}$	49.642 ± 2.5131

Table 4.2: Relative error for voltages and angles using IEEE 39-bus system and likelihood-free methods.

As we commented before, ABC, K2ABC, PABC obtained similar results. However, we observe that the posterior means for angles computed by PABC are the closest to the values calculated by MCS. Similarly, we note that the results for the posterior mean in voltages obtained by ABC are the closest to the values using MCS. Notice that the ABC MCMC and ABC SMC computed poor results for the the posterior means in both cases. For the standard deviations, we observe that the results obtained for angles and voltages do not coincide to the values computed by MCS, the errors computed by these methods were lower than 26% and 68% for angles and voltages, respectively.

Finally, notice that despite assuming an informative prior for $\mathbf{x}$, the posterior distributions obtained by ABC, K2ABC, PABC, ABC MCMC and ABC SMC are not close to the densities obtained by MCS. In the next subsections, we will focus on the PPF analysis using the ABC, JABC, ABC SMC and JABC SMC. We will also apply directly these methods over the observed and simulated data, since we need additional computational steps to obtain the summary statistics.

4.2.2 PPF analysis with different sample sizes and number of nodes

In this experiment, we are interested in observing the performance of the different methods when the sample size for the input random variables and the number of the nodes of

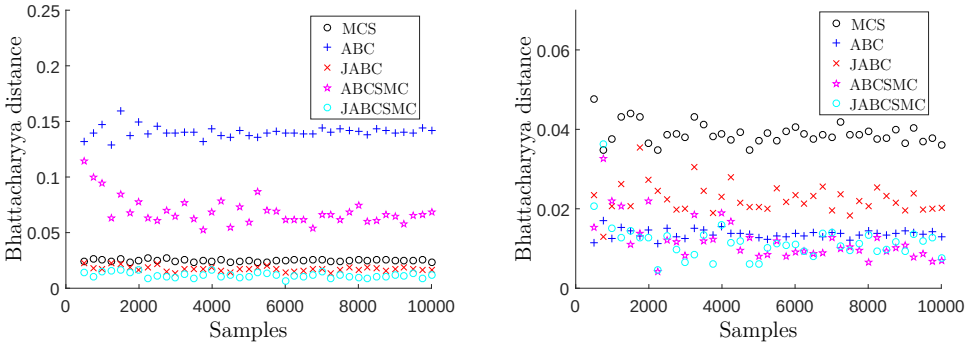
the PS were increased. Therefore, we considered two case studies: in the first case, we analyzed the PPF problem for the IEEE 6-bus system assuming that the true posterior of $\mathbf{x}$ was also known and then observed how the simulation methods worked with different sample sizes. For the second case, we examined classic PPF problems in terms of the sample and system size. For classic PPF problems, we generated N_s samples from the input random variables to obtain $\{\mathbf{b}_i\}_{i=1}^{N_s}$ from Eq. (2.1). We then inferred the posterior $p(\mathbf{x}|\,\mathbf{b})$ using all methods over the IEEE $\{6, 39, 118\}$ bus systems.

For the first case, we used the input random variables shown in Su (2005) and we increased the number of samples of these variables from 500 to 10000 samples, in steps of 250 samples. For each step, we ran ABC, JABC, ABC SMC JABC SMC, MCS and computed the Bhattacharyya distance (BD) (Kailath, 1967) to measure the similarity between the true distribution over $\mathbf{x}$ and the distribution computed by each method. For ABC and JABC, we used $\epsilon = 0.7$. For ABC SMC and JABC SMC, we used $\{\epsilon_t\}_{t=1}^{T} = \{3.0, 2.0, 1.0, 0.9, 0.7\}$. From now on, we will continue to employ Gaussian and uniform distributions as prior distributions for each voltage and angle, respectively, as it was described in the first experiment of the previous subsection. For ABC SMC, we employed a multivariate Gaussian distribution as proposal distribution for $\mathbf{x}$, with a covariance matrix that depends on $\sigma_{v_q}^2 = 10^{-6}$ and $\sigma_{\theta_q}^2 = 10^{-7}$ as diagonal entries for voltages and angles, respectively.

For JABC and JABC SMC, we employed a multivariate Gaussian distribution as symmetric proposal distribution, $q \sim \mathcal{N}(\mathbf{x}_i, \mathbf{\Sigma}_{\mathbf{x}})$, where $\mathbf{x}_i$ and $\mathbf{\Sigma}_{\mathbf{x}}$ are the mean and covariance matrix for the proposal distribution. The mean $\mathbf{x}_i$ can be computed using Eq. (4.8). The covariance $\mathbf{\Sigma}_{\mathbf{x}}$ is assumed to be a diagonal matrix with parameters, $\sigma_{v_q}^2 = 10^{-5}$ and $\sigma_{\theta_q}^2 = 10^{-6}$. For $\mathbf{x}_0$, we used a vector of ones for voltages, and the DC power flow solution for angles.

In Fig. 4.3, we present the BD in terms of the number of samples. We show the BD for voltages (see Fig. 4.3 (left)) and angles (see Fig. 4.3 (right)). From Fig. 4.3 (left), we observe that the BD obtained by MCS is the most steady. We also see that JABC and JABC SMC compute better BD than MCS, ABC and ABC SMC; being the BD computed by JABC SMC, the lowest. For angles, we note that the BD calculated by ABC is the most steady, but the BD achieved by JABC SMC is the lowest. BD values are summarized in the Table 4.3. These results validate our hypothesis of using the

Jacobian of the power flow equations into the ABC algorithm to search more probable samples for $\mathbf{x}$. As can be seen in Fig. 4.3, the JABC SMC is a proper estimator of the posterior over $\mathbf{x}$ in terms of the number of samples. From Fig. 4.3, ABC and ABC SMC computed better BDs for angles than the BDs for voltages, this is due to the information included by DC power flow solution in the prior over angles.



BD for voltages BD for angles

Figure 4.3: BD for voltages and angles for different number of samples. The circles, plus signs, cross, stars and the squares represent the BD obtained by MCS, ABC, JABC, ABCMCS and JABCSMC, respectively. *Left:* show the BD for voltages. *Right:* show the BD for angles.

Method	BD_V	BD_θ
MCS	0.0248 ± 0.0009	0.0389 ± 0.0027
ABC	0.1404 ± 0.0049	0.0136 ± 0.0011
ABCSMC	0.0690 ± 0.0126	0.0121 ± 0.0054
JABC	0.0174 ± 0.0022	0.0225 ± 0.0037
JABCSMC	$\mathbf{0.0117 \pm 0.0024}$	$\mathbf{0.0117 \pm 0.0051}$

Table 4.3: BD obtained by all ABC methods and MCS.

For the second case study, we solved two classic PPF problems with different number of nodes. In particular, we analyzed the IEEE 6, 39 and 118 bus systems using input Gaussian variables (Su, 2005). We initially applied all ABC methods and MCS to these three systems to infer the posterior $p(\mathbf{x}|\,\mathbf{b})$ over $N_s = 1000$ samples from the input random variables. We used the parameters mentioned in the previous experiment for the IEEE 6 bus test system. In the IEEE 39 bus test system, we chosen ϵ of 9 and 2 for ABC and JABC, respectively. the sequence of thresholds for ABC SMC are chosen as $\{\epsilon_t\}_{t=1}^{T} = (100, 50, 25, 20, 15, 10, 9)$. Similarly for JABC SMC, we employ $\{\epsilon_t\}_{t=1}^{T} =$

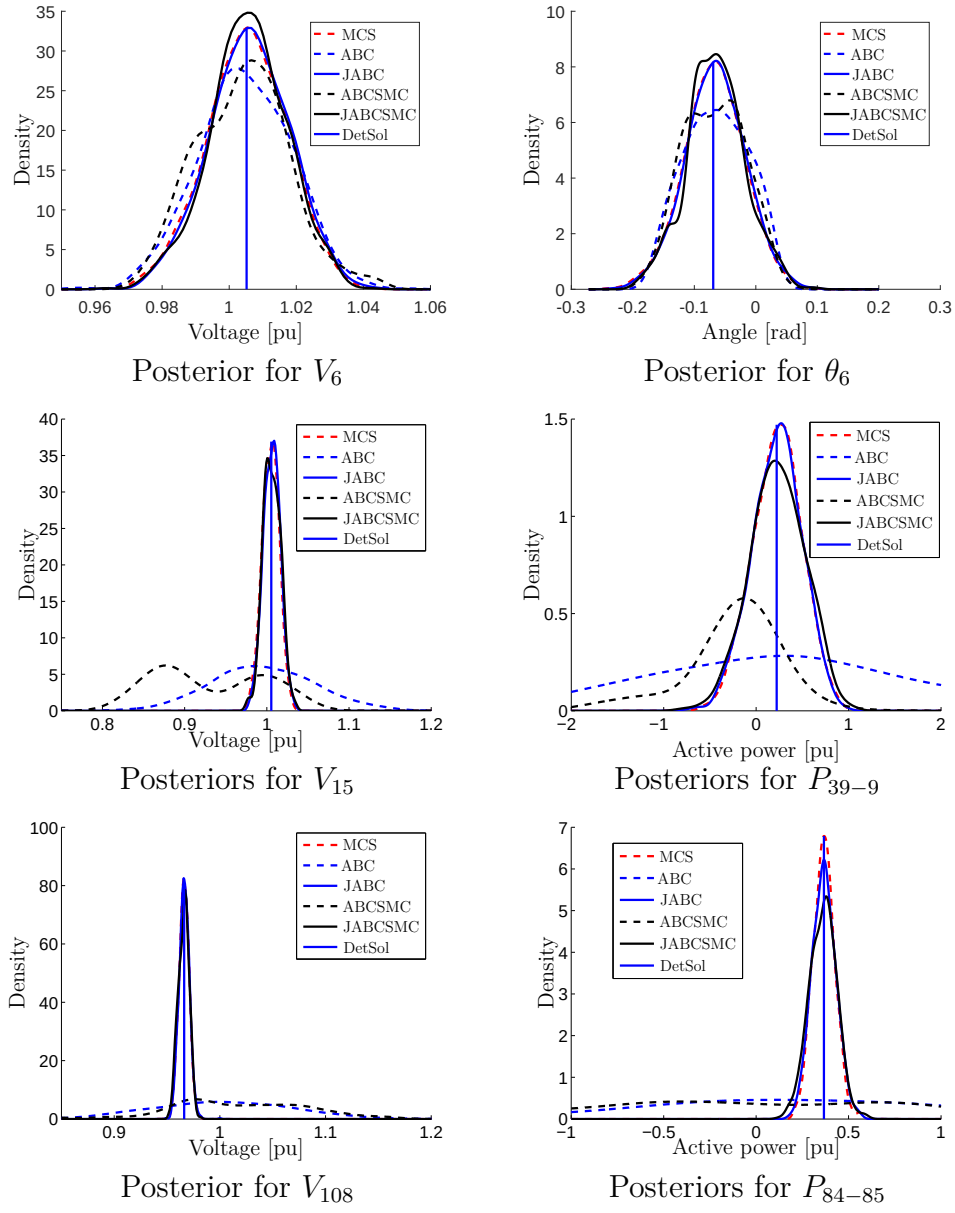
$(10, 9, 8, 7, 6, 5, 4, 3, 2)$. For ABC SMC, JABC and JABC SMC, we used $\sigma^2_{v_q} = 10^{-4}$ and $\sigma^2_{\theta_q} = 10^{-5}$ as parameters of the proposal distributions. For the IEEE 118 bus test system, we used $\epsilon = 4.0$ in the ABC and JABC algorithms. For ABC SMC and JABC SMC, we used $\{\epsilon_t\}_{t=1}^{T} = \{10.0, 9.0, 8.0, 7.0, 6.0, 5.0, 4.0\}$. For ABC SMC, JABC and JABC SMC, we employed $\sigma^2_{v_q} = 10^{-5}$ and $\sigma^2_{\theta_q} = 10^{-6}$ as parameters of the proposal distributions. Fig. 4.4 compares the MCS and all ABC methods when some variables of these systems are inferred. Specifically, we show the probability distribution over V_6 and θ_6 (see Fig. 4.4 in top row) for the 6 bus system, we also present the posterior for V_{15} and P_{39-9} (see Fig. 4.4 in middle row) using the IEEE 39 bus test system, and the distribution for θ_{108} and active power between nodes 84 and 85 (P_{84-85}) (see Fig. 4.4 in bottom row) for the 118 bus system, since the results for other variables in $\mathbf{x}$ are similar to the results shown in Fig. 4.4. [5] From Fig. 4.4 in top row, we observe that ABC and ABC SMC acceptably estimate of the posterior of V_6 and θ_6. For V_{15}, P_{39-9}, V_{108} and P_{84-85}, ABC and ABC SMC obtain accepted samples around the DPF solution (see the blue vertical solid line), however the shape of the posteriors obtained by these methods is not similar to the one obtained using MCS (see Fig. 4.4 in middle and bottom rows). On the other hand, JABC and JABC SMC provide results close to the distributions obtained by MCS, confirming the importance of the improved state space exploration of $\mathbf{x}$ in the Jacobian ABC-type methods.

We also recorded the computation time (CT) required for each method to solve the PPF problems shown in Fig. 4.4. Table 4.4 lists the CTs took for the different methods. Notice that JABC took the lowest CT for solving the PPF problems in both systems. Notice also that ABC and ABC SMC approaches are the least efficient methods for solving the PPF problem. ABC took the highest CT to analyze the uncertainty in both systems. This high CT is due to the low acceptance rates to ensure the matching between $\mathcal{D}'$ and $\mathcal{D}$. MCS also took high CTs for the two systems.

Considering input Gaussian variables, we compared the DPF solution and the estimated posterior mean using the relative error (RE), [6] for each variable using the different solution methods. In what follows, the numerical analysis is only done for the state variables (state variables can be used to compute other variables, i.e. active and reactive

[5] After obtaining samples from the posterior of $\mathbf{x}$, we employ the Eq. (2.2) for computing the posterior over P_{39-9} and P_{84-85}.

[6] The RE is computed between the DPF solution and the estimated posterior mean obtained by each method, see the appendix B for more details.



Posterior for V_6 Posterior for θ_6

Posteriors for V_{15} Posteriors for P_{39-9}

Posterior for V_{108} Posteriors for P_{84-85}

Figure 4.4: Posterior for V_6, θ_6, V_{15}, P_{39-9}, θ_{108} and P_{84-85}. The dashed red, blue and black lines are the responses obtained by MCS, ABC and ABC SMC. The solid blue and black lines are the posteriors calculated by JABC and JABC SMC. The blue vertical solid line is the DPF solution. *Top row:* show the posteriors for V_6 and θ_6 using the IEEE 6 bus test system. *Middle row:* show the posteriors for V_{15} and P_{39-9} using the IEEE 39 bus test system. *Bottom row:* show the posteriors for V_{108} and P_{84-85} using the IEEE 118 bus test system.

Computation time [s]	MCS	JABC	JABC SMC	ABC	ABC SMC
IEEE-6	13.540	**1.0140**	4.3056	640.11	9.2664
IEEE-39	15.100	**1.2636**	11.512	938.04	26.145
IEEE-118	17.097	**4.1496**	32.198	1277.2	52.428

Table 4.4: Computation time obtained by MCS, JABC, JABC SMC, ABC and ABC SMC, in the solution of the PPF problem using the IEEE 6-bus, 39-bus and 118-bus test system.

power flows between buses). For the comparison, we drew 25 subsets of input random variables, we then applied all ABC methods and MCS to obtain 25 REs for angles and voltages. With these errors, we computed the mean and the standard deviation for each method. We used the $\{6, 39, 118\}$-bus systems and 1000 samples for all input random variables. We slightly changed the application of ABC methods. Here we ran 1000 simulations for the ABC methods by each input random variable configuration, since in the previous experiments, we used one simulation by each $\mathbf{b}_i$. Due to the ABC methods are approximate inference approaches, we wanted to show if it was possible to obtain REs close to the results using MCS. Table 4.5 presents the RE obtained by MCS and ABC methods.

Index [%]	methods	IEEE 6	IEEE 39	IEEE 118
RE_θ	MCS	2.5413 ± 1.6234	2.4465 ± 1.6187	4.7911 ± 2.2914
	ABC	8.9756 ± 0.1614	12.832 ± 0.0095	58.008 ± 0.0220
	ABC SMC	9.2201 ± 0.1847	12.882 ± 0.0851	98.398 ± 0.1031
	JABC	$\mathbf{2.5282 \pm 1.6163}$	$\mathbf{2.2512 \pm 1.7008}$	$\mathbf{2.3266 \pm 1.1796}$
	JABC SMC	2.5378 ± 1.6205	2.2518 ± 1.6936	2.3267 ± 1.1794
RE_v	MCS	0.0571 ± 0.0194	$\mathbf{0.0383 \pm 0.0181}$	$\mathbf{0.0226 \pm 0.0044}$
	ABC	0.2661 ± 0.0150	2.4951 ± 0.0015	2.3997 ± 0.0008
	ABC SMC	0.6683 ± 0.0095	2.9346 ± 0.0167	2.4896 ± 0.0036
	JABC	$\mathbf{0.0559 \pm 0.0192}$	0.0413 ± 0.0119	0.0239 ± 0.0027
	JABC SMC	0.0569 ± 0.0193	0.0398 ± 0.0114	0.0240 ± 0.0027

Table 4.5: Relative errors between the deterministic solution of the systems (see the vertical lines of the above figures) and the estimated posterior mean for angles and voltages using all methods when all input random variables are modeled by Gaussian distributions.

For the IEEE 6-bus test system, notice that the REs obtained by JABC and JABC SMC are lower than the REs computed by MCS, ABC and ABC SMC for voltages

and angles. For the IEEE 39 and 118-bus test systems, we note that the RE for angles calculated by JABC and JABC SMC are lower than the one obtained by MCS, ABC and ABC SMC. We also note that MCS computed the lowest RE for the voltages, however the REs from JABC and JABC SMC are close to results achieved by MCS. Finally, notice that ABC and ABC SMC obtained adequate results for the IEEE 6-bus test system, the errors computed by these two methods were lower than 10% and 1% for angles and voltages, respectively. However for large systems, the errors obtained by ABC and ABC SMC were lower than 3.0% for voltages, and they were greater than 10% for angles. These results show that it is necessary (for the ABC and ABC SMC) to use an informative prior over $\mathbf{x}$ as it was presented in the results shown in Fig. 4.2.

After showing the results when we had 1000 samples from the input random variables, we repeated the experiment of Table 4.5 with 100, 500, 1000 and 2000 samples. For this experiment, we only used the MCS, JABC and JABC SMC. Figs 4.5 and 4.6 report the RE, for angles and voltages, versus the number of samples. Blue, green and brown bars are REs obtained by MCS, JABC and JABC SMC, respectively. The vertical line represents one standard deviation for each RE.

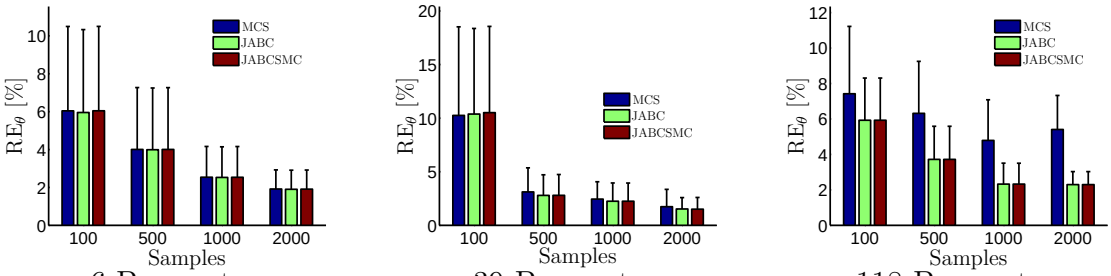
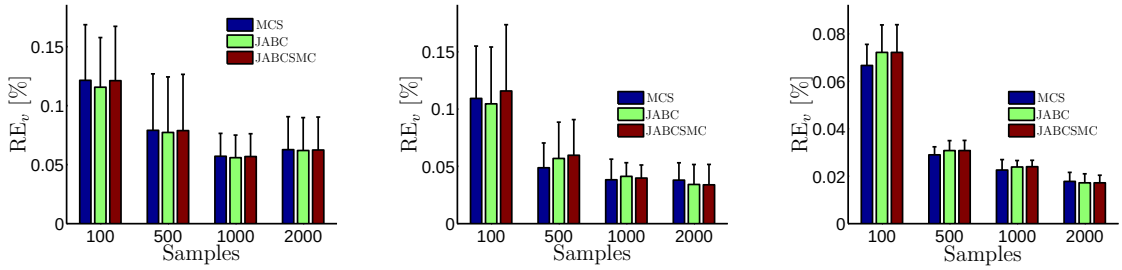


Figure 4.5: Relative errors (RE) for angles when the number of the samples from the input random variables is increased. Blue bars are the RE obtained by MCS, green bars are the RE for JABC method and the brown bars are the RE using JABC SMC. Vertical lines represent one standard deviation for each variable. *Left:* Relative errors for the IEEE 6-bus system. *Center:* Relative errors for the IEEE 39-bus system. *Right:* Relative errors for the IEEE 118-bus system.

Figs 4.5 and 4.6 show that the RE decreases when the number of samples increases in the three systems. From these two figures, the REs obtained by JABC and JABC SMC tend to REs computed by MCS. However from Fig. 4.5 (right), the REs obtained by JABC and JABC SMC are lower than the REs using MCS. These high REs obtained



6 Bus system 39 Bus system 118 Bus system

Figure 4.6: Relative errors (RE) for voltages when the number of the samples from the input random variables is increased. Blue bars are the RE obtained by MCS, green bars are the RE for JABC approach and the brown bars are the RE using JABC SMC. Vertical lines represent one standard deviation for each variable. *Left:* Relative errors for the IEEE 6-bus system. *Center:* Relative errors for the IEEE 39-bus system. *Right:* Relative errors for the IEEE 118-bus system.

by MCS are due to poor estimates for the angle at node 26, causing high variability (see the standard deviation in each scenario). We also note that for the IEEE 39-bus system using 2000 samples, the REs for voltages are lower than the REs for these variables shown in table 4.5 where we used 1000 samples. These results confirm that the ABC methods, which are approximate Bayesian inference approaches, provide satisfactory results with respect to the optimization-based simulation methods.

4.2.3 PPF analysis with other methodologies

In this subsection, we are interested to compare our Jacobian ABC methods against other methodologies, for example: the point-estimation method based on two points (PEM) (Su, 2005) and the Taguchi method (TM) (Hong et al., 2016), to analyze PPF problems. For this comparison, we solved two classic PPF problems with different number of nodes. In particular, we analyzed the IEEE 6 and 118 bus test systems using input Gaussian variables (Su, 2005). We applied JABC, JABC SMC and also used MCS to these two systems to infer the posterior $p(\mathbf{x}|\,\mathbf{b})$ over $N_s = 1000$ samples from the input random variables. For the PEM, due to this method must be combined with some series expansion to acquire the probability distribution of the PPF results, we used a Gram-Charlier series expansion (Morales and Perez-Ruiz, 2007). For the TM, we employed a nonparametric density estimator (Bowman and Azzalini, 1997) to obtain the probability distributions of the PPF results. We used the parameters mentioned in the previous experiment for both systems.

Fig. 4.7 compares each method against the MCS when some variables of these systems are inferred. Specifically, we show the probability distribution over V_6 (see Fig. 4.7 in the top row) for the 6 bus system and the distribution for θ_{108} (see Fig. 4.7 in the bottom row) for the 118 bus system, since the results for other variables in $\mathbf{x}$ are similar to the results shown in Fig. 4.7. From Fig. 4.7 in the first column, we observe that the PEM does not infer appropriately the probability distributions of both variables. We also note that the samples obtained by the TM in both variables (see Figs 4.7 in the second column) are around the DPF solution (see the blue vertical solid line), however the shapes of the probability distributions obtained by this method are not similar to the probability distributions obtained by MCS. However, JABC and JABC SMC provide results close to the distributions obtained by MCS (see Figs. 4.7 in the third and fourth columns), confirming the importance of the improved state space exploration of $\mathbf{x}$ in the Jacobian ABC-type methods.

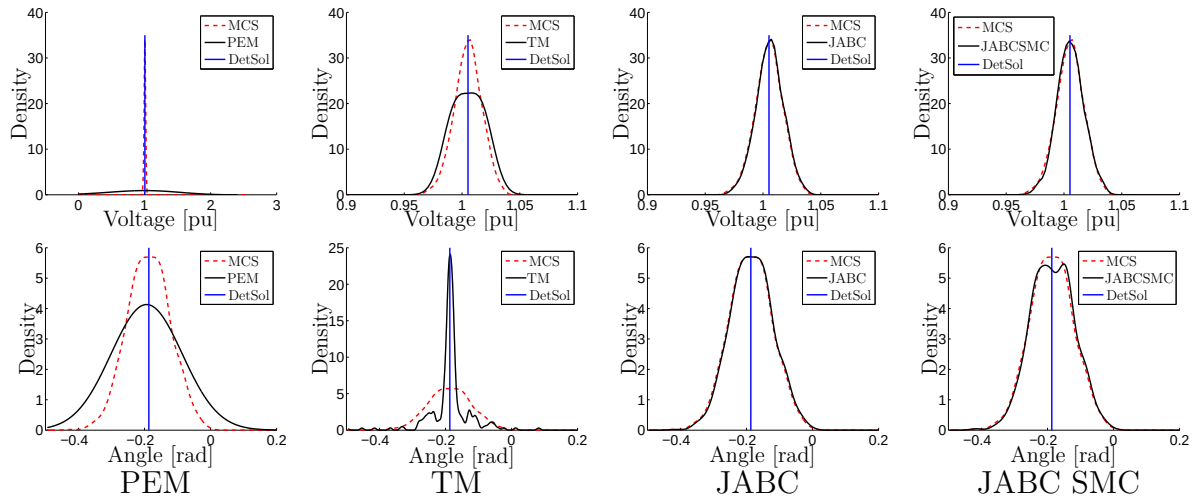


Figure 4.7: Posterior for V_6 and θ_{108}. The dashed red line is the responses obtained by MCS. The solid black lines are the probability densities calculated by PEM, TM, JABC and JABC SMC. The blue vertical solid line is the DPF solution. *Top row:* Posterior for V_6 using the IEEE 6 bus test system. *Bottom row:* Posterior for θ_{108} employing the IEEE 118 bus test system.

Table 4.6 lists the CTs took for the different methods to solve the PPF problems shown in Fig. 4.7. Notice that the CTs, for JABC and JABC SMC, are lower than the CT took by MCS. From Table 4.6, we also note that the proposed methods require

more CT compared to the PEM and TM. However, the proposed methods do not require DPF solutions to obtain the probability densities shown in Fig 4.7. It is necessary to comment that the goal of the PEM and TM is to reduce of the number of simulations in MCS, therefore the number of samples is also reduced. From this point of view, a fair comparison between these two methods and the remaining methods applied would not be possible.

		MCS	PEM	TM	JABC	JABC SMC
CT [s]	IEEE-6	13.540	0.1248	0.0936	0.6365	1.7156
	IEEE-118	17.097	6.0528	5.2884	5.8464	16.757

Table 4.6: Computation time (CT), in seconds [s], required by MCS, PEM, TM, JABC and JABC SMC methods to solve two PPF problems when assuming input Gaussian variables.

Using input Gaussian variables, we compared the performance of different solution methods through of the relative error (see Eq. (B.1)) between the DPF solution and the estimated posterior mean for each variable. Similarly to the previous experiments, we drew 25 subsets of input random variables, we then applied PEM, TM, JABC, JABC SMC and MCS to obtain 25 REs for angles and voltages. Finally, we computed the mean value and an one standard deviation for each method. For this comparison, we used the $\{6, 118\}$-bus systems and 1000 samples for all input random variables. Table 4.7 lists the relative using the different methods mentioned before.

Index	Method	IEEE 6	IEEE 118
RE_θ [%]	MCS	1.5190 ± 0.8843	3.7290 ± 1.9822
	PEM	1.5268 ± 0.8923	3.7841 ± 2.0324
	TM	1.5494 ± 0.9092	3.7674 ± 2.0281
	JABC	$\mathbf{1.5073 \pm 0.8793}$	2.3344 ± 1.0952
	JABC SMC	1.5149 ± 0.8813	$\mathbf{2.3341 \pm 1.0951}$
RE_V [%]	MCS	0.0598 ± 0.0302	$\mathbf{0.0212 \pm 0.0054}$
	PEM	0.0603 ± 0.0300	0.0213 ± 0.0055
	TM	0.0605 ± 0.0300	0.0212 ± 0.0054
	JABC	$\mathbf{0.0585 \pm 0.0302}$	0.0234 ± 0.0019
	JABC SMC	0.0596 ± 0.0302	0.0235 ± 0.0019

Table 4.7: Relative error (RE) for angles and voltages using all methods when using input Gaussian variables.

For the 6-bus system, notice that the REs obtained by JABC and JABC SMC are lower than the REs computed by MCS, PEM and TM for voltages and angles. For the

118-bus system, the REs computed by JABC and JABC SMC are lower than the results calculated by MCS, PEM and TM for angles. We also note that MCS and TM computed the lowest RE for the voltages, however the REs from JABC and JABC SMC are close to the results achieved by MCS.

4.2.4 PPF analysis with renewable energy

Next, we proceeded to consider renewable energy in a IEEE 39-bus test system. Similar to the experiment presented in the subsection 3.3.4, we model all input random variables as in Soleimanpour and Mohammadi (2013), but we do not consider correlated loads. We also add a wind farm at bus 39.

In this experiment, we drew 1000 samples from the input random variables and we used the same prior distributions over $\mathbf{x}$ as it was mentioned in the first experiment. However, we increased the variance in the Gaussian prior for each voltage to $\sigma_v^2 = 0.005$. We only compared the MCS, JABC and JABC SMC. We used $\epsilon = 2.0$ in the JABC algorithm. For JABC SMC, we used $\{\epsilon_t\}_{t=1}^{T} = \{3.0, 2.75, 2.5, 2.25, 2.0\}$. For JABC and JABC SMC, we employed $\sigma_{v_q}^2 = 10^{-5}$ and $\sigma_{\theta_q}^2 = 10^{-6}$ as parameters of the proposal distributions. We have chosen the same variables shown in Fig. 4.8.

Fig. 3.5 compares the posterior distributions for θ_{12}, θ_{28}, θ_{34}, θ_{39}, V_1, V_4, V_{14} and V_{24} obtained by MCS, JABC and JABC SMC. From this figure, we notice that JABC inferred satisfactory the posterior distributions for the angles. We also observe that JABC SMC estimated acceptably the posterior distributions for the angles, however, it is possible to note that the posterior means and the standard deviation obtained for each angle are not similar to the results obtained by MCS.

On the other hand, from Fig. 4.8, we observe that JABC and JABC SMC inferred properly the probability densities for voltages. Despite being approximate inference methodologies, the JABC and JABC SMC, computed results close to those obtained by MCS. MCS, JABC and JABC SMC took 34.491s, 2.5428s and 19.780s, respectively, for solving this PPF problem. These results show how the likelihood-free Bayesian simulators is an alternative for solving the PPF problems, allowing a non-linear relationship between $\mathbf{b}$ and $\mathbf{x}$, and dealing with input non-Gaussian variables. We also analyzed the results for the remaining variables, and they are similar to the results shown in Fig. 4.8.

Since we do not have a ground-truth solution, we observed how close the mean and

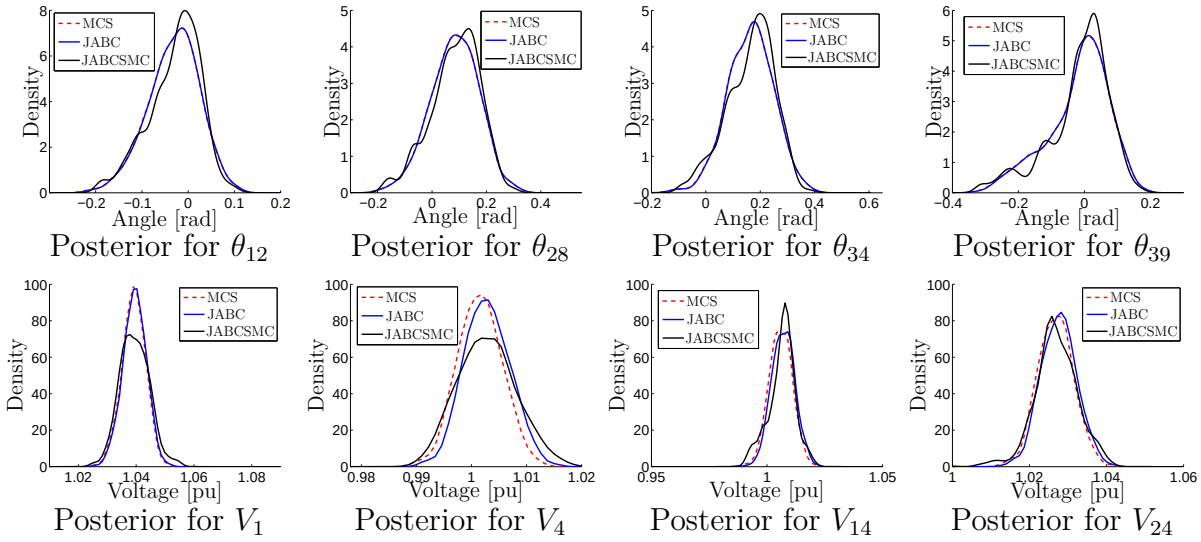


Figure 4.8: Posterior for some state variables. The dashed red line is the responses obtained by MCS. The solid blue and black lines are the probability densities calculated by JABC and JABC SMC. *Top row:* Posterior for θ_{12}, θ_{28}, θ_{24} and θ_{39}. *Bottom row:* Posterior for V_1, V_4, V_{14} and V_{24}.

standard deviation of angles and voltages obtained by the ABC methods are to the values computed by MCS, that is, we used the mean and standard deviation obtained by MCS as reference values. Using these values, we computed the relative error for ABC methods. We generated 25 different sets of input random variables, with 1000 samples for each variable. We then applied the JABC and JABC SMC to infer the posterior mean and standard deviation of each input random variable configuration. Finally, we computed 25 relative errors using the previous information. Table 4.8 lists the mean and one standard deviation for the relative errors that compare the reference values obtained by MCS, and the estimated values using the ABC methods. Notice that JABC gives better estimated results than JABC SMC, since in some cases the estimates obtained from JABC SMC have a large variance. However, the errors obtained by JABC SMC do not exceed 3%.

IEEE 39		$\varepsilon_\theta[\%]$	$\varepsilon_V[\%]$
JABC	μ	0.6866 ± 0.0768	0.1251 ± 0.0026
	σ	0.2017 ± 0.1254	0.3751 ± 0.1300
JABC SMC	μ	2.6610 ± 1.8072	0.1173 ± 0.0375
	σ	2.0547 ± 1.6072	2.3299 ± 1.0785

Table 4.8: Relative error for voltages and angles using IEEE 39-bus system.

4.2.5 Discussion

Several solution methods for the PPF problem do not take into account the fact that previous knowledge of state variables might be available in terms of probability distributions (Soleimanpour and Mohammadi, 2013; Su, 2005). In this study we demonstrated that it is possible to perform this aspect in the PPF analysis. By including prior distributions over the state variables into the PPF problem, we can include additional information for tackling PPF problems.

We must mention that it is important (for the ABC and ABC SMC) to use an improved prior distribution for $\mathbf{x}$, since if we use a non-informative prior over $\mathbf{x}$, the ABC and ABC SMC can only work for small power systems using input Gaussian variables. We also show how to choose the summary statistics for ABC methods, however their performance also depends on how to select the prior distribution over $\mathbf{x}$.

From the results obtained using the large test systems and the ABC methods (ABC, K2ABC, PABC, ABC MCMC and ABC SMC), we demonstrate, despite assuming an informative prior for $\mathbf{x}$, that it is necessary the inclusion of the Jacobian matrix information to these methods to do an improved state space exploration.

When we use the Jacobian of the power flow equations as part of the ABC methods, we observe that the Jacobian ABC methods are able to obtain probability densities close to the probability distributions using MCS, that is, these results validate our hypothesis of using the Jacobian of the power flow equations into the ABC algorithm to search more probable samples for $\mathbf{x}$. However, it is necessary to choose adequately the initial condition $\mathbf{x}_0$; otherwise, the Jacobian ABC methods could obtain samples with low probability.

Several solution methods have been proposed to address the PPF problem. We have compared the performance of our Jacobian ABC methods against two of these proposed methods, such as the PEM and TG. These methods can be computationally more effective than our methods and MCS, however they must be combined with mathematical approaches to obtain the probability densities of the PPF solutions (Ren et al., 2016). Our methods are competitive alternatives for solving PPF problems with respect to MCS, and besides, they consider the state variables as random variables.

We have also provided access to the MATLAB code used to reproduce some experiments, which is publicly available at `https://github.com/cardazuluaga/PPFdemos.git`

Chapter 5

Conclusions and Future work

This chapter presents the conclusions about work done in this thesis, and it also considers some future lines of research.

5.1 Conclusions

The goal of this thesis was to propose Bayesian approaches for tackling PPF problems.

Likelihood-based methods for PPF analysis. In Chapter 3 a hierarchical Bayesian model for the probabilistic power flow analysis was proposed. In this chapter, we specified prior distributions over the state variables and a likelihood function that relates $\mathbf{x}$ and $\mathcal{D}$. This Bayesian PPF model was inferred by using two likelihood-based methods (MCMC and HMC) and two different test systems. We demonstrated that the use of this model and the HMC provide better results than those obtained by MCMC.

Likelihood-free methods for PPF analysis. We introduced an alternative for solving PPF problems using the ABC method and the Jacobian of the power flow equations. We also proposed priors for voltages and angles for the PPF problem under a Bayesian inference perspective. We demonstrated that ABC and ABC SMC can work for an small power system using input Gaussian variables. However, it is necessary to define an informative prior over the state variables. We also demonstrated the use of Kernel Embeddings as summary statistics for ABC methods, however their performance also depends on how to select the prior distribution over $\mathbf{x}$. We also showed that the posteriors of the state variables obtained by JABC and JABC SMC are close to the results using MCS, similarly JABC took less computation time for obtaining the PPF

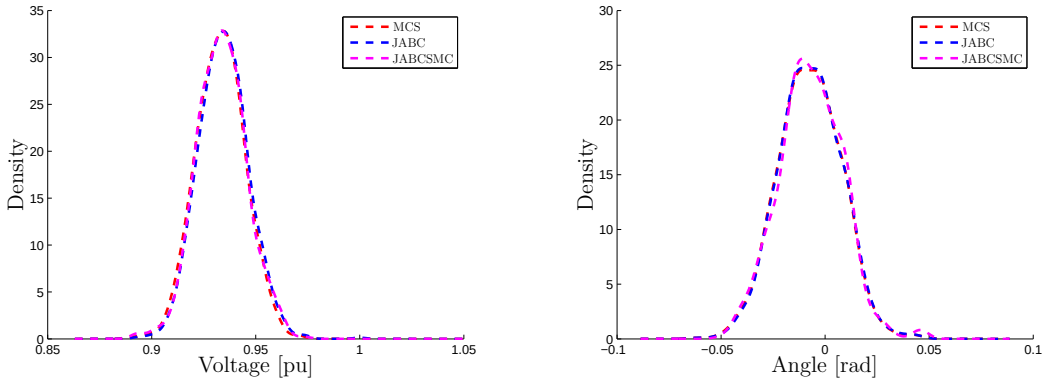
solution with respect to MCS.

On the other hand, we validated our methods using different test systems and compared the performance of these methods against three methods, such as MCS, the PEM and TG. We demonstrated that our approaches consider uncertainty in state variables in the PPF analysis and are alternatives for solving PPF problems.

5.2 Potential research lines

As future works, it would be possible to consider:

- In chapters 3 and 4, we proposed Bayesian approaches using independent random variables and we do not consider uncertainty over the line parameters. We do believe that it is possible to contemplate these two aspect into a Bayesian approach.
- In this study, we focused in formalizing Bayesian inference for PPF analysis over transmission systems as power systems. These systems are widely analyzed, however we observe that it would be interesting the application of our method to other areas in power systems, for example, in distribution networks.



Posterior for V_{28} Posterior for θ_{16}

Figure 5.1: Posterior for θ_{16} and V_{28} considering a IEEE 33-bus radial distribution system. The dashed red, blue and magenta lines are the responses obtained by MCS, JABC and JABC SMC, respectively. *Left:* Posterior for V_{28}. *Right:* Posterior for θ_{16}.

As an initial example, we have applied the Jacobian ABC methods to analyze the PPF problem for the IEEE 33-bus radial distribution system,[1] assuming that we have

[1]This system is available at `http://www.pserc.cornell.edu/matpower/`

input Gaussian variables. We specifically applied MCS, JABC and JABC SMC. We generated 1000 samples from the input random variables to obtain $\{\mathbf{b}_i\}_{i=1}^{1000}$. We then inferred the posterior $p(\mathbf{x}|\,\mathbf{b})$ using these three methods. In the Fig. 5.1, we compare the posterior distribution for V_{28} (left) and θ_{16} (right) when MCS, JABC and JABC SMC are applied. We observe that the JABC and JABC SMC provide similar results to the distributions obtained by MCS.

We have included a proof of concept of the application of our methods to analyze a radial network, showing promising results in a single-phase balanced radial configuration. However, it is necessary to analyze this system over different conditions, for example: to consider unbalanced radial systems, distribution systems with large amounts of wind and solar generation, or distribution systems approximated by radial or convex models.

Appendix A

IEEE test systems

A.1 IEEE 6-bus test system

For our experiments, we use the IEEE 6-bus test system that includes 3 load demands, 11 transmission lines and 3 generators as shown in Fig. A.1. Table A.1 shows the transmission line data. In Table A.2, the parameters for the conventional generators are shown. Then in Table A.3, we present the parameters for the loads.

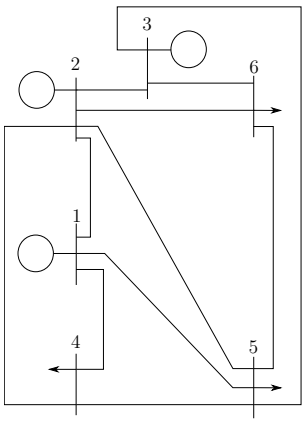


Figure A.1: IEEE 6 bus test system (Su, 2005). Circles represent electric generators; vertical bars are the nodes of the system; lines between two nodes represent the electrical connection between two nodes (transmission lines); arrows are loads.

From bus	To bus	r_{km} (p.u.)	x_{km} (p.u.)	$b_{km}/2$ (p.u.)
1	2	0.1	0.2	0.02
1	4	0.05	0.2	0.02
1	5	0.08	0.3	0.03
2	3	0.05	0.25	0.03
2	4	0.05	0.1	0.01
2	5	0.1	0.3	0.02
2	6	0.07	0.2	0.025
3	5	0.12	0.26	0.025
3	6	0.02	0.1	0.01
4	5	0.2	0.4	0.04
6	6	0.1	0.3	0.03

Table A.1: Transmission line data for the system shown in Fig. A.1.

Bus	1	2	3
type	*Slack*	*Gen.*	*Gen.*
Voltage (p.u.)	1.05	1.05	1.07
Active Power (p.u.)	–	0.5	0.6

Table A.2: Parameters for the generators.

Bus	P_d (p.u.)	Q_d (p.u.)
4	0.7	0.7
5	0.7	0.7
6	0.5	0.7

Table A.3: Parameters for the load. P_d y Q_d are the active and reactive power, respectively.

A.2 IEEE 39 and 118-bus test systems

As part of our experiments, we also use the IEEE 39-bus and IEEE 118-bus test systems. For the IEEE 39-bus test system, which has 10 generation units, 46 transmission lines, and 29 load demands. The IEEE 118-bus test systems has 53 generation units, 186 transmission lines, and 99 load demands. [1]

[1] These systems are available at `http://www.pserc.cornell.edu/matpower/`

Appendix B

Validation metrics

For validation purposes, we compute the relative error (RE) between the value obtain using DLF analysis without considering uncertainty over variables and the mean value for each variable obtained using the simulation methods. RE is given by

$$RE = \left| \frac{x_{wu} - x_{mp}}{x_{wu}} \right|, \tag{B.1}$$

where x_{wu} is the original value without considering uncertainty over variables, and x_{mp} is the most probable value for that variable, obtained through each simulation method. This error is only used when the real and reactive power at PQ nodes, ad voltages and real power injected at PV nodes are modeled by Gaussian distributions. For the wind energy is considered, we compute relative errors with respect to the values calculated using MCS, that is, we employ (Su, 2005),

$$\varepsilon_x^{\mu} = \left| \frac{\mu_x - \mu_x^*}{\mu_x} \right|, \tag{B.2}$$

$$\varepsilon_x^{\sigma} = \left| \frac{\sigma_x - \sigma_x^*}{\sigma_x} \right|, \tag{B.3}$$

where ε_x^{μ} and ε_x^{σ} are relative error for the mean and standard deviation values; μ_x and σ_x are the mean and standard deviation obtained using MCS; μ_x^* and σ_x^* are the mean and standard deviation computed with Bayesian simulators.

We also use the Bhattacharyya distance (BD) for measuring the similarity between the true distribution over the state variables and the distribution computed by each method (Kailath, 1967).

Appendix C

Publications

C.1 Published Papers

- C. Zuluaga, M. Álvarez, E. Giraldo, Short-term wind speed prediction based on robust Kalman filtering: An experimental comparison, Applied Energy, Vol. 156, pp. 321 – 330, 2015.

- C. Zuluaga, E. Valencia, M. Álvarez, A. Orozco, A Parzen-Based Distance Between Probability Measures as an Alternative of Summary Statistics in Approximate Bayesian Computation, Lecture Notes in Computer Science, vol 9279, pp. 50 – 61, 2015.

- C. Zuluaga, M. Álvarez, Approximate Probabilistic Power Flow, In: Woon W., Aung Z., Kramer O., Madnick S. (eds) 4th International Workshop on Data Analytics for Renewable Energy Integration. DARE 2016. Lecture Notes in Computer Science, vol. 10097, pp. 43 – 53, 2016.

- C. Zuluaga, M. Álvarez, Bayesian Probabilistic Power Flow Analysis Using Jacobian Approximate Bayesian Computation, to appear at IEEE Transactions on Power Systems.

References

Aien, M., Fotuhi-Firuzabad, M., and Rashidinejad, M. (2014). Probabilistic optimal power flow in correlated hybrid wind –photovoltaic power systems. *Smart Grid, IEEE Transactions on*, 5(1):130–138. (pages 2, 8, 9, 13, 14, 16, and 17)

Allan, R., Borkowska, B., and Grigg, C. (1974). Probabilistic analysis of power flows. *Electrical Engineers, Proceedings of the Institution of*, 121(12):1551–1556. (page 8)

Allan, R. and Leite da Silva, A. (1981). Probabilistic load flow using multilinearisations. *Generation, Transmission and Distribution, IEE Proceedings C*, 128(5):280–287. (page 9)

Allan, R., Leite da Silva, A., and Burchett, R. (1981). Evaluation methods and accuracy in probabilistic load flow solutions. *Power Apparatus and Systems, IEEE Transactions on*, PAS-100(5):2539–2546. (page 9)

Allan, R. N. and Al-Shakarchi, M. R. G. (1976). Probabilistic a.c. load flow. *Electrical Engineers, Proceedings of the Institution of*, 123(6):531–536. (page 8)

Allan, R. N. and Al-Shakarchi, M. R. G. (1977). Probabilistic techniques in a.c. load-flow analysis. *Electrical Engineers, Proceedings of the Institution of*, 124(2):154–160. (page 8)

Berlinet, A. and Christine, T.-A. (2004). *Reproducing Kernel Hilbert Spaces in Probability and Statistics*. Springer Science+Business Media, LLC. (page 33)

Bishop, C. M. (2006). *Pattern Recognition and Machine Learning*. Springer. (pages 2, 11, 13, and 18)

Blum, M., Nunes, M., Prangle, D., and Sisson, S. (2012). A comparative review of dimension reduction methods in Approximate Bayesian Computation. arXiv:1202.3819. (page 33)

Blum, M. G. (2010). Choosing the summary statistics and the acceptance rate in Approximate Bayesian Computation. In Lechevallier, Y. and Saporta, G., editors, *Proceedings of COMPSTAT'2010*, pages 47–56. Physica-Verlag HD. (page 33)

Blum, M. G. B., Nunes, M. A., Prangle, D., and Sisson, S. A. (2013). A comparative review of dimension reduction methods in Approximate Bayesian Computation. *Stat. Sc.*, 28(2):189–208. (page 33)

Borkowska, B. (1974). Probabilistic load flow. *Power Apparatus and Systems, IEEE Transactions on*, PAS-93(3):752–759. (page 8)

Bowman, A. W. and Azzalini, A. (1997). *Applied smoothing techniques for data analysis : the kernel approach with S-Plus illustrations*. Clarendon Press ; Oxford University Press. (page 50)

Bracale, A., Carpinelli, G., Di Fazio, A. R., and Russo, A. (2013). *Probabilistic Approaches for the Steady-State Analysis of Distribution Systems with Wind Farms*, pages 245–282. Springer Berlin Heidelberg, Berlin, Heidelberg. (page 37)

Briceno, W., Caire, R., and Hadjsaid, N. (2012). Probabilistic load flow for voltage assessment in radial systems with wind power. *International Journal of Electrical Power and Energy Systems*, 41(1):27 – 33. (page 9)

Cai, D., Shi, D., and Chen, J. (2014). Probabilistic load flow with correlated input random variables using uniform design sampling. *International Journal of Electrical Power and Energy Systems*, 63:105 – 112. (page 9)

Carmona-Delgado, C., Romero-Ramos, E., and Riquelme-Santos, J. (2015). Probabilistic load flow with versatile non-Gaussian power injections. *Electric Power Systems Research*, 119:266 – 277. (pages 1, 8, 9, and 11)

Chen, T., Fox, E. B., and Guestrin, C. (2014). Stochastic Gradient Hamiltonian Monte Carlo. In *Proceedings of the 31st International Conference on Machine Learning, Cycle 2*, volume 32 of *JMLR Proceedings*, pages 1683–1691. JMLR.org. (page 20)

Ding, N., Fang, Y., Babbush, R., Chen, C., Skeel, R. D., and Neven, H. (2014). Bayesian sampling using stochastic gradient thermostats. In Ghahramani, Z., Welling, M., Cortes, C., Lawrence, N., and Weinberger, K., editors, *Advances in Neural Information Processing Systems 27*, pages 3203–3211. Curran Associates, Inc. (page 20)

Dong, Z., Wang, Y., and Zhao, J. (2014). Variational Bayesian inference for the probabilistic model of power load. *IET Generation, Transmission Distribution*, 8(11):1860–1868. (page 11)

Dopazo, J. F., Klitin, O. A., and Sasson, A. M. (1975). Stochastic load flows. *IEEE Transactions on Power Apparatus and Systems*, 94(2):299–309. (pages 2, 8, 9, 12, and 40)

Fan, M., Vittal, V., Heydt, G., and Ayyanar, R. (2013). Probabilistic power flow analysis with generation dispatch including photovoltaic resources. *Power Systems, IEEE Transactions on*, 28(2):1797–1805. (page 10)

Fearnhead, P. and Prangle, D. (2012). Constructing summary statistics for Approximate Bayesian Computation: semi-automatic Approximate Bayesian Computation. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 74(3):419–474. (page 33)

Gallego, L. A. and Padilha-Feltrin, A. (2012). Power flow for primary distribution networks considering uncertainty in demand and user connection. *International Journal of Electrical Power and Energy Systems*, 43(1):1171 – 1178. (pages 1, 9, and 10)

Gelman, A., Carlin, J., Stern, H., Dunson, D., Vehtari, A., and Rubin, D. (2013). *Bayesian Data Analysis, Third Edition*. Chapman & Hall/CRC Texts in Statistical Science. Taylor & Francis. (page 30)

Grainger, J. and Stevenson, W. (1994). *Power system analysis*. McGraw-Hill series in electrical and computer engineering: Power and energy. McGraw-Hill. (pages 7 and 37)

Gretton, A., Borgwardt, K. M., Rash, M. J., Schölkopf, B., and Smola, A. (2007). A kernel method for the two-sample problem. In *Advances in Neural Information Processing Systems 15*, pages 513–520. MIT Press. (page 34)

Gretton, A., Borgwardt, K. M., Rash, M. J., Schölkopf, B., and Smola, A. (2012). A Kernel Two-Sample Test. *Journal of Machine Learning Research*, 13:723–773. (page 34)

Hong, Y. Y., Lin, F. J., and Yu, T. H. (2016). Taguchi method-based probabilistic load flow studies considering uncertain renewables and loads. *IET Renewable Power Generation*, 10(2):221–227. (pages 10, 38, and 50)

Huelsenbeck, J. P., Ronquist, F., Nielsen, R., and Bollback, J. P. (2001). Bayesian inference of phylogeny and its impact on evolutionary biology. *Science*, 294(5550):2310–2314. (page 4)

Jiang, B., Wu, T.-Y., Zheng, C., and Wong, W. H. (2015). Learning summary statistic for Approximate Bayesian Computation via deep neural network. *arXiv preprint arXiv:1510.02175*. (page 33)

Kailath, T. (1967). The divergence and Bhattacharyya distance measures in signal selection. *IEEE Transactions on Communication Technology*, 15(1):52–60. (pages 44 and 63)

Kalesar, B. and Seifi, A. R. (2010). Fuzzy load flow in balanced and unbalanced radial distribution systems incorporating composite load model. *International Journal of Electrical Power and Energy Systems*, 32(1):17 – 23. (page 10)

Le, D., Berizzi, A., Bovo, C., Ciapessoni, E., Cirio, D., Pitto, A., and Gross, G. (2013). A probabilistic approach to power system security assessment under uncertainty. In *Bulk Power System Dynamics and Control - IX Optimization, Security and Control of the Emerging Power Grid (IREP), 2013 IREP Symposium*, pages 1–7. (pages 1, 9, and 10)

Leite da Silva, A. and Arienti, V. L. (1990). Probabilistic load flow by a multilinear simulation algorithm. *Generation, Transmission and Distribution, IEE Proceedings C*, 137(4):276–282. (page 9)

Marjoram, P., Molitor, J., Plagnol, V., and Tavaré, S. (2003). Markov Chain Monte Carlo without Likelihoods. *Proceedings of the National Academy of Sciences of the United States of America*, 100(26):15324–15328. (page 36)

Meeds, T. and Welling, M. (2015). Optimization Monte Carlo: Efficient and embarrassingly parallel likelihood-free inference. In *Advances in Neural Information Processing Systems 28*, pages 2080–2088. Curran Associates, Inc. (page 37)

Mohamed, S., Heller, K. A., and Ghahramani, Z. (2012). Bayesian and l1 approaches for sparse unsupervised learning. In *Proceedings of the 29th International Coference on International Conference on Machine Learning*, ICML'12, pages 683–690, USA. Omnipress. (page 14)

Morales, J. and Perez-Ruiz, J. (2007). Point estimate schemes to solve the probabilistic power flow. *Power Systems, IEEE Transactions on*, 22(4):1594–1601. (pages 10, 22, and 50)

Morales, J. M., Baringo, L., Conejo, A. J., and Minguez, R. (2010). Probabilistic power flow with correlated wind sources. *IET Generation, Transmission Distribution*, 4(5):641–651. (page 1)

Mori, H. and Jiang, W. (2009). A new probabilistic load flow method using MCMC in consideration of nodal load correlation. In *Intelligent System Applications to Power Systems, 2009. ISAP '09. 15th International Conference on*, pages 1–6. (page 11)

Murphy, K. P. (2012). *Machine Learning: A Probabilistic Perspective (Adaptive Computation and Machine Learning series)*. The MIT Press. (pages 2, 11, 16, 18, 19, and 36)

Nunes, M. A. and Balding, D. J. (2010). On optimal selection of summary statistics for Approximate Bayesian Computation. *Statistical Applications in Genetics and Molecular Biology*, 9(1). (page 33)

Park, M., Jitkrittum, W., and Sedjdinovic, D. (2015). K2-ABC: Approximate Bayesian Computation with infinite dimensional summary statistics via kernel embeddings. arXiv:1502.02558. (pages 33, 34, 35, and 41)

Parry, K. and Hazelton, M. L. (2013). Bayesian inference for day-to-day dynamic traffic models. *Transportation Research Part B: Methodological*, 50:104 – 115. (page 4)

Prusty, B. R. and Jena, D. (2017). A critical review on probabilistic load flow studies in uncertainty constrained power systems with photovoltaic generation and a new approach. *Renewable and Sustainable Energy Reviews*, 69:1286 – 1302. (pages 8 and 9)

Raiffa, H. and Schlaifer, R. (2000). *Applied Statistical Decision Theory*. Wiley Classics Library. Wiley. (page 11)

Ren, Z., Li, W., Billinton, R., and Yan, W. (2016). Probabilistic power flow analysis based on the stochastic response surface method. *IEEE Transactions on Power Systems*, 31(3):2307–2315. (page 55)

Saunders, C. (2013). Point estimate method addressing correlated wind power for probabilistic optimal power flow. *Power Systems, IEEE Transactions on*, PP(99):1–10. (page 10)

Shimazaki, H. and Shinomoto, S. (2010). Kernel bandwidth optimization in spike rate estimation. *J. Comput. Neurosci.*, 29(1-2):171–182. (page 42)

Sisson, S. A., Fan, Y., and Tanaka, M. M. (2007). Sequential Monte Carlo without likelihoods. *Proceedings of the National Academy of Sciences*, 104(6):1760–1765. (page 37)

Soleimanpour, N. and Mohammadi, M. (2013). Probabilistic load flow by using nonparametric density estimators. *Power Systems, IEEE Transactions on*, 28(4):3747–3755. (pages 1, 2, 8, 9, 10, 15, 22, 24, 27, 53, and 55)

Sriperumbudur, B., Gretton, A., Fukumizu, K., Schölkopf, B., and Lanckriet, G. (2010). Hilbert space embeddings and metrics on probability measures. *Journal of Machine Learning Research*, 11:1517–1561. (page 33)

Su, C. (2005). Probabilistic load-flow computation using point estimate method. *Power Systems, IEEE Transactions on*, 20(4):1843–1851. (pages 7, 14, 21, 22, 24, 25, 38, 39, 44, 45, 50, 55, 60, and 63)

Sun, L., Lu, Y., Jin, J., Lee, D.-H., and Axhausen, K. (2015). An integrated Bayesian approach for passenger flow assignment in metro networks. *Transportation Research Part C: Emerging Technologies*, 52:116 – 131. (page 4)

Tebaldi, C. and West, M. (1998). Bayesian inference on network traffic using link count data. *Journal of the American Statistical Association*, 93(442):557–573. (page 4)

Toni, T., Welch, D., Strelkowa, N., Ipsen, A., and Stumpf, M. (2009). Approximate Bayesian Computation scheme for parameter inference and model selection in dynamical systems. *Journal of the Royal Society Interface*, 6:187–202. (pages 31, 36, and 38)

Usaola, J. (2009). Probabilistic load flow with wind production uncertainty using cumulants and Cornish - Fisher expansion. *International Journal of Electrical Power and Energy Systems*, 31(9):474–481. (page 10)

Usaola, J. (2010). Probabilistic load flow with correlated wind power injections. *Electric Power Systems Research*, 80(5):528 – 536. (page 10)

Valverde, G., Saric, A. T., and Terzija, V. (2012). Probabilistic load flow with non-Gaussian correlated random variables using Gaussian mixture models. *IET Generation, Transmission Distribution*, 6(7):701–709. (page 11)

Vorsic, J., Muzek, V., and Skerbinek, G. (1991). Stochastic load flow analysis. In *[1991 Proceedings] 6th Mediterranean Electrotechnical Conference*, pages 1445–1448 vol.2. (pages 8 and 9)

Wang, Z. and Alvarado, F. (1992). Interval arithmetic in power flow analysis. (page 9)

Wilkinson, D. J. (2007). Bayesian methods in bioinformatics and computational systems biology. *Briefings in Bioinformatics*, 8(2):109–116. (page 4)

Wilkinson, R. (2013). Approximate Bayesian Computation (ABC) gives exact results under the assumption of model error. *Statistical Applications in Genetics and Molecular Biology*, 12(2):129–141. (pages 31 and 32)

Wood, A. and Wollenberg, B. (1996). *Power Generation, Operation, and Control*. A Wiley-Interscience publication. Wiley. (page 10)

Wu, L., Shahidehpour, M., and Li, T. (2007). Stochastic security-constrained unit commitment. *IEEE Transactions on Power Systems*, 22(2):800–811. (page 14)

Yu, H., Chung, C., Wong, K. P., Lee, H., and Zhang, J. (2009). Probabilistic load flow evaluation with hybrid latin hypercube sampling and Cholesky decomposition. *Power Systems, IEEE Transactions on*, 24(2):661–667. (page 9)

Zhang, H. and Li, P. (2010). Probabilistic analysis for optimal power flow under uncertainty. *Generation, Transmission Distribution, IET*, 4(5):553–561. (pages 1, 9, and 10)

Zhang, P. and Lee, S. (2004). Probabilistic load flow computation using the method of combined cumulants and Gram-Charlier expansion. *Power Systems, IEEE Transactions on*, 19(1):676–682. (page 10)

Zhu, J. (2015). *Optimization of power system operation (ieee press series on power engineering); 2nd ed.* Wiley-Blackwell. (pages 8, 23, and 40)